

The cost of silence?

Analyses of AIDS deaths attributable to misguided policies in South Africa carry lessons for scientific leaders.

It is not often that the cost of political leaders being wrong gets quantified. It's doubly important to bring science to bear on such a question when people have paid with their lives, as happened with the AIDS epidemic in South Africa under the leadership of then-president Thabo Mbeki.

An article from the Harvard School of Public Health AIDS Initiative published this week (P. Chigwedere *et al.* *J. Acq. Immun. Def. Synd.* **49**, 410–415; 2008) provides such a service by estimating the benefits lost through underuse of antiretroviral (ARV) drugs in South Africa. Using modelling and clinical data of ARV-drug efficacy, it compares the number of people who received ARV therapy and of pregnant women who received treatment to prevent mother-to-child transmission of HIV between 2000 and 2005 with the number who could feasibly have been treated during the same period. It concludes that the lack of an ARV drug programme caused the loss of more than 330,000 lives — consistent with an estimate along different lines by the South African economist Nicoli Nattrass (N. Nattrass *Afr. Affairs* **107**, 157–176; 2008) — and that 35,000 babies were needlessly born with HIV. The authors, by their own account, have been conservative in deriving these estimates.

Mbeki was deposed in September, and his successor, Kgalema Motlanthe, moved swiftly to replace Mbeki's chief accomplice, health minister Manto Tshabalala-Msimang, with Barbara Hogan. Hogan has rejected the dissident line, subscribed to by both her predecessor and Mbeki, that AIDS is not caused by HIV. Moreover, describing herself as “ashamed” about the estimates, she has declared that “the era of denialism is over completely in South Africa”.

This is almost certainly true. Motlanthe, who supported Mbeki's stand on HIV-AIDS at the time, has his support base in the country's trade-union movement, which was never happy about Mbeki's line on this issue. Jacob Zuma, leader of the African National Congress and a presidential contender in next year's election, also supported Mbeki's standpoint before he was fired by him. He now espouses only populist causes, and has moved on from denialism about AIDS: the South African electorate has no more appetite for it.

The needless deaths that occurred in South Africa prompt reflection

on Mbeki's now infamous presidential AIDS advisory panel on the link between HIV and AIDS, the fate of which was chronicled in this journal in 2000–01. Its inconclusive report enabled Mbeki and his cabinet, who must bear collective responsibility, to portray this link as “deeply contested, and contestable”, to quote Nattrass. Certainly, the AIDS dissidents (much criticized by *Nature* in the past) couldn't wait to participate in the panel. But should orthodox scientists have signed up?

Even in retrospect, this is a difficult question to answer. Once leading South African scientists, such as Malegapuru Makgoba, then president of the South African Medical Research Council and an outspoken critic of Mbeki, had agreed to do so, others were bound to follow suit in support. In turn, members from outside of the country in good faith believed that their colleagues deserved similar support, and so agreed to participate. Ultimately it became clear that these efforts were a waste of time, as there was no possibility of consensus being reached among the panel's two diametrically opposed camps.

Mbeki used this lack of consensus to justify a national policy that refrained from rolling out ARV drugs until late 2003, although some of the country's nine provinces, which enjoy a level of autonomy on matters relating to health, defied this with varying degrees of success. Even as late as 2006 the South African health department's stand at the International AIDS Conference was extolling the virtues of garlic, beetroot and lemon juice as a solution to the epidemic.

In retrospect, the panel, constituted as it was, should never have been supported. Yet several of the country's key scientific institutions explicitly endorsed its establishment, and also desisted from criticizing Mbeki. Along with his cabinet, they bear some culpability for the consequences that have now been documented. There is a moral to this tragic tale that may prove relevant in other contexts. In a young democracy with a historically hierarchical culture, and with attitudes often hardened by a colonial past, scientific institutions need not only to guard their independence fiercely but also to make their reasoned voices heard above the fray of political sycophancy. ■

“The South African electorate has no more appetite for AIDS denialism.”

Culture clash in China

An online row highlights the need for Chinese universities to fix their hiring policies.

In October, an online war broke out between Yi Rao, a neuroscientist and Peking University's dean of life sciences, and Keming Cui, a plant biologist and professor emeritus at the university who has a string of positions on academic and awards committees and editorial posts on Chinese journals to his name.

Cui retired from Peking University four years ago but kept his laboratory there. This year he tried to have his associate professor formally take over the lab. Such transfers of power are common in China, but they are also criticized as a way for powerful professors to hold lab space beyond their tenure. Rao refused to acknowledge the transfer of authority. Instead he planned to drastically cut the lab's size.

Rao, the first Peking University dean to be hired through an international search, says he wants to ensure that the university hires the best faculty members through appropriate evaluation. He also wants to ensure that qualified outsiders are considered and that an inbred

academic system is avoided. Rao says that the associate professor will have a few years to prove himself before he is evaluated for promotion and to see whether he can keep the laboratory.

On 9 October, Cui began writing a string of entries in his blog, which became widely read when copied by other websites and the online bulletin boards of Peking University. Cui described Rao's action as belittling his field of plant anatomy because it was not a 'hot' area. He made a stand for basic science. The blog drew some sympathetic comments from students, who copied it to more widely read student blogs. Rao, whose own blog normally gets about 2,000 hits per entry, immediately posted his defence, which picked up 10,000 hits.

Newspapers hesitate to pick up such hot potatoes, so the debate devolved to the blogs. Although they offer a platform for such discussion, blogs also make irresponsible name-calling possible. After alleging that Rao was trying to cut off support for a discipline of science, Cui compared Rao's efforts to the activities of Trofim Lysenko — the Soviet 'state scientist' who in the 1940s used his close connections to the Soviet leadership to crush scientists who opposed his views.

Both scientists are in all likelihood doing what is natural to them

to promote the next generation of scientists. The situation is further confused by Peking University's lack of clear guidelines on how to proceed in such situations, leaving new regulations set against old customs. When Rao took over as dean in September 2007, the university made it clear in writing that such hiring decisions would be his to make.

Deans and university presidents in China are watching to see how the situation is resolved. Will Cui, who worked at Peking University for more than 40 years, be able to raise public support and use his connections with senior colleagues to get his way? Or will Rao be able to stand his ground? It should, and looks as if it will, be the latter.

But further changes are needed. China and its universities now have the money to undertake proper recruitment exercises, and more universities should be seeking candidates beyond their walls and outside China's borders. They need clear, consistently applied guidelines on who has the authority to make those decisions. Of course, although a clear policy would be good, vesting that much power in one individual's hands, as Rao recognizes, requires caution. So regulations that check and make transparent the actions of those given decision-making power will also be needed. ■

Focus on Earth

Europe is rightly pioneering the systematic appliance of science in space to societal needs.

To achieve unanimity among European countries is no mean feat, particularly when it involves spending billions. So congratulations are due to the ministers of the 18 member states of the European Space Agency (ESA) who meet every three years to agree on funds, and who briskly struck a deal at their summit in the Netherlands last week giving the agency its requested €10 billion (US\$12.8 billion), which includes €2.34 billion in new commitments (see page 552).

The ministers also delivered a well-merited vote of confidence in ESA's science programme, which has long punched well above its financial weight, and has enjoyed a string of prominent successes. It got every euro it asked for — which was hardly a given considering the present economic backdrop.

That happy outcome fits within a broader shift in which the European Union (EU) seems gradually to be adopting a coherent twenty-first-century space policy. Indeed, it is getting its act together in a far more focused and visionary fashion than the United States, which nonetheless boasts much larger space budgets. The EU's major new focus is to deliver tangible benefits of space activity to citizens and society, and to address key challenges such as climate change and natural disasters, with Earth monitoring as its flagship. Importantly, fundamental research has not been sacrificed, and is set to thrive.

The planned joint EU–ESA Kopernikus constellation of Earth observation satellites, scheduled for launch over the next decade, will deliver a wealth of real-time data and maps of planet Earth at our keyboard fingertips. This represents a far-sighted vision for the

use of space science to meet societal needs. And, whereas in the past research satellites were often one-offs, Kopernikus will translate the expertise and technologies acquired from generations of research satellites into fleets of operational satellites delivering data 24/7, year-in, year-out.

The focus of Kopernikus on generating long-term continuous data sets, and exploiting these data as user-friendly services is one that European space science could benefit and learn from. One of ESA's weaknesses, as highlighted recently by the European Science Foundation is that, unlike NASA, it has no clear remit to fund archiving and support for scientific analyses of all the space data it collects. So this falls largely to national agencies, whose role here the foundation describes as "inadequate in volume, fragmented, and dictated by national concerns".

ESA ministers took steps to remedy this last week by approving a €72.3-million initiative to mine past and future data for essential climate variables defined by scientists. But much greater progress could be made, and Europe needs to pay more attention to making better use of the flood of data streaming in from above.

Making data more widely available and accessible, in particular to support policy, is a cornerstone for the EU, which over the past decade has assumed political leadership of space policy in Europe. This shift in power, away from ESA and national governments, was confirmed in September in a resolution that gave the EU the lead role in coordinating investment and operation of Europe's space activities.

The respective roles of the EU and ESA in both Kopernikus and the global navigation system Galileo point to an evolution in which ESA will become the autonomous science and technology development arm of the EU, and the EU will take responsibility for the heavy-lifting of costly operational application and infrastructure. Such change is desirable, with ESA being the place where innovative technologies are developed and where inspiring science is done — in turn, driving the applications of tomorrow. ■

RESEARCH HIGHLIGHTS

ASTRONOMY

Blast from the past

Astrophys. J. doi:10.1086/589761 (2008)

The famous supernova 1987A has a cousin. In 2001, astronomers at Pennsylvania State University working with NASA's Chandra X-ray Observatory spotted stellar fireworks in the Circinus galaxy, which is 13 million light years from Earth — dose in astronomical terms. The blast seemed like a supernova, but, at the time, Franz Bauer could not confidently say what had set the sparks off.

A team of astronomers led by Bauer of Columbia University in New York has now confirmed that the X-ray source is a type II supernova, dubbed SN 1996cr, light from which first reached Earth a dozen years ago. The blast got steadily brighter in terms of X-rays and radio waves; both of these types of radiation were a thousand times brighter than in SN 1987A's case.

CLIMATE CHANGE

Mulled wine

Climate Res. doi:10.3354/cr00759 (2008)

During the past few decades the Priorat region of northeastern Spain has risen from oenocultural near-abandonment to producing some of the world's most coveted wines. But climate change may make sustaining the pre-eminence of these cult wines ever harder.

María Concepción Ramos at the University of Lleida in Spain and her colleagues analysed climate records from 1952 to 2006. They found that climate change — especially increasing maximum temperature — is pushing the Priorat climate into a range that may be too warm to support the intense yet balanced wines for which the region is known. In an area with severe water limitations, the decreases in early-season precipitation that the team recorded may prove even more challenging than temperature change.



Silent killer

Science 322, 1392–1395 (2008)

Experimental drugs called bicyclic nitroimidazoles kill non-replicating as well as replicating bacteria, making them promising candidates for the treatment of latent tuberculosis. One such compound, PA-824, does this by a mechanism that suggests a new avenue for anti-infective drug design, report Clifton Barry of the US National Institute of Allergy and Infectious Diseases in Bethesda, Maryland, and his colleagues.

PA-824 inhibits cell division by interfering with the manufacture of bacteria's sticky outer layers. Barry and his colleagues report that it also generates nitric oxide, which is toxic regardless of whether or not a cell is dividing. PA-824 should not have this effect on mammalian cells or on all bacteria — only on bacteria such as *Mycobacterium tuberculosis* that produce a compound known as F₄₂₀. This is needed for the drug to trigger nitric oxide release.



D. GUERCHOIS/REUTERS

NEUROBIOLOGY

Mad mouse disease

Neuron 60, 598–609 (2008)

Studies of Creutzfeldt–Jakob disease (CJD), of which 'mad cow' disease is a form, have suffered from a lack of suitable animal models. A newly developed mutant mouse could reinvigorate the field.

To create it, Roberto Chiesa of the Mario Negri Institute for Pharmacological Research in Milan, Italy, and his colleagues expressed a disease-associated variant of prion protein (PrP) in mice. The mutant mice exhibited all the classic symptoms of CJD, including a lack of coordination, accumulation of PrP aggregates in the brain, memory loss and abnormal electroencephalographic and sleep–wake patterns. Previous models of CJD have replicated only the motor effects of the disease.

depends on the rank of the variable according to a power law. Now Thomas Maillart and his colleagues at ETH Zurich in Switzerland report empirical evidence from the spread of open-source software that an explanation posited in 1955 is correct.

This came from economist Herbert Simon. He thought that Zipf's law stems from the growth of a population of which the size varies at random but with a standard deviation proportional to that size. An analysis of exceptionally detailed data reveals that the Zipf law in incoming links to packages of the Debian Linux computer-operating system is supported by exactly this growth process.

EVOLUTION

Electric love

Biol. Lett. doi:10.1098/rsbl.2008.0566 (2008)

Fish from the lower reaches of the Congo River, of the genus *Campylomormyrus*, use the discharge of their electric organ to find prey. Such discharges may also be driving species divergence, suggest experiments by Philine Feulner, now of the University of Sheffield, UK, and her colleagues.

They gave ready-to-spawn

STATISTICS

One size fits all

Phys. Rev. Lett. 101, 218701 (2008)

In the 1940s, linguist George Kingsley Zipf found that the probability distribution of a wide range of variables, including word-use frequency and demographic distributions,

C. compressirostris females a choice between males of the same species, and either a closely related but electrically different *C. rhynchophorus*, or a distantly related but electrically similar *C. tamandua*. The procedure was repeated with a simulation of the males' electrical discharges in place of the male fishes.

The electric signal seems to be all-important. In both tests, females preferred their own kind over *C. rhynchophorus*, but not over *C. tamandua*.

PHARMACOLOGY

Setting the pace

Cell Metab. **8**, 482–491 (2008)

A protein targeted by some diabetes drugs might also help to regulate daily cycles in blood pressure and heart rate.

Thiazolidinediones are widely prescribed for the treatment of type 2 diabetes, and work by activating a protein called PPAR- γ . Tianxin Yang of the University of Utah in Salt Lake City and his colleagues investigated the cardiovascular role of this protein using mice that lacked expression of the gene that encodes PPAR- γ in vascular smooth-muscle cells.

Rhythmic variations in blood pressure and heart rate, as well as in the expression of several body-clock genes in blood vessels, were diminished in these mice. Furthermore, normal mice treated with a thiazolidinedione called rosiglitazone showed increased expression of *Bmal1*, a dock gene.

QUANTUM PHYSICS

Signature shift

Science **322**, 1357–1360 (2008)

The discovery of the Lamb shift in 1947 was a key factor in the development of quantum electrodynamics (QED). This slight discrepancy between the predicted and observed energy levels of electrons in a hydrogen atom was eventually explained by QED as the effect of virtual photons flickering in and out of existence in a vacuum. Now that effect has been replicated in an electrical circuit by Andreas Wallraff at ETH Zurich in Switzerland and his colleagues.

Their work uses 'artificial atoms', quantum devices designed to have an energy spectrum resembling that of real atoms. Tuning a coupling between such artificial atoms and a microwave transmission line containing a strong vacuum allowed them to produce a Lamb shift of 1.4% for a typical emission line.

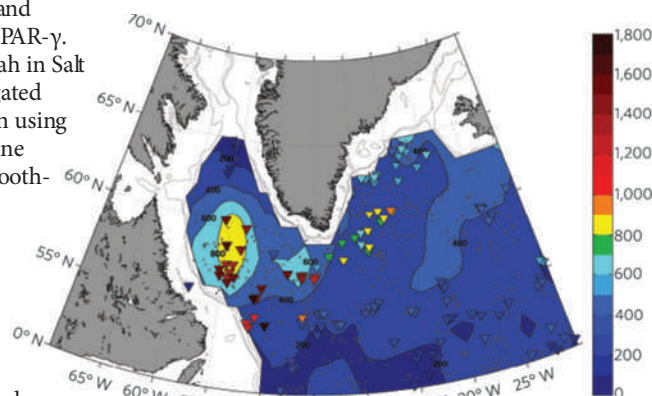
GENOMICS

The baby-milk bacterium

Proc. Natl Acad. Sci. USA **105**, 18964–18969 (2008)

The guts of breast-fed babies often contain bacteria of the subspecies *Bifidobacterium longum infantis*, which keep at bay harmful microorganisms that might otherwise take up residence. The genome of this subspecies has now been published.

Of its 2,423 recognized protein-coding genes, 702 are not found in related bacterial taxa sequenced so far, write David Mills of the University of California, Davis, and his colleagues. The bacterium feeds on specific sugars in human milk that infants do not use, and Mills and his team have identified several genes that make this possible. Another gene cluster makes enzymes that allow the microbes to consume milk-borne urea in the nitrogen-poor environment of the infant bowel.



GEOSCIENCES

Deep-sea mix

Nature Geosci. doi:10.1038/ngeo382 (2008)

After more than a decade of shallow or non-existent convection currents in the Labrador Sea — with only a brief return to deep-water convection at the turn of the century — the carbon-dioxide-sequestering subpolar gyre in the North Atlantic Ocean seems to have bounced back.

Kjetil Våge of the Woods Hole Oceanographic Institution in Massachusetts and his colleagues report data from the system of floats known as the Argo programme. These show a return last winter to mixing as deep as 1,800 metres in the Labrador Sea, 1,000 metres in the Irminger Sea and 1,600 metres south of Greenland. Triangles in the image (above) indicate these depths (in metres).

The change happened surprisingly quickly. The reasons for it include unusually low air temperatures, an increased flux of fresh water and pack-ice from the Arctic and changes in the North Atlantic storm track.

JOURNAL CLUB

John Greally
Albert Einstein College of
Medicine, New York City

A biologist considers a link between jumping genes and immune-system enzymes.

Many viruses present a fierce threat to the body. They contain nucleic acids that, when free to roam in a cell's cytoplasm, elicit an immune response involving proteins called interferons. Pairings of the nucleic-acid residues cytosine and guanine are especially good at this, unless they carry a chemical modification in the form of a methyl group. This modification is the norm for 'jumping genes', or retrotransposons, which can move around the human genome and were probably once viral genes themselves.

A team led by Daniel Stetson at the University of Washington in Seattle has uncovered a useful twist to this tale. While searching for proteins that interact with cytoplasmic nucleic acids, the researchers came across Trex1. Mutated versions of Trex1 are known to cause chilblain lupus in humans, and in mice lead to autoimmune myocarditis, whereby the immune system attacks the heart. Stetson *et al.* say that mice lacking Trex1 have huge numbers of retrotransposons in their heart muscles.

Critically, the authors' molecular surveys reveal that Trex1 suppresses the rate at which jumping genes move around. This indicates that Trex1 protects the body from misidentifying its own parts as 'foreign' by degrading retrotransposons and thus preventing them from overloading the system (D. B. Stetson *et al. Cell* **134**, 587–598; 2008).

That jumping genes have the potential to overwhelm the system in this way was unexpected. Most experts had assumed that the addition of methyl groups took care of quenching them. But if retrotransposons are made at a rate that triggers inflammation, as Stetson and his colleagues' experiments propose, it could open up a whole new avenue for research. Everyone studying lupus and related diseases should be excited.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS

CHINA PHOTOS/GETTY IMAGES



Chemical reagents can take months to arrive in China.

Slow shipping hobbles Chinese science

A lack of laboratory reagents is taking its toll on researchers.

SHANGHAI

Earlier this year, a researcher in Shanghai requested antibodies from a collaborator at Harvard University. Working with customs and shippers to arrange the transfer, including extra dry ice to keep the materials cool, cost US\$1,500 and took well over a month. "It was very tedious and very expensive," says the scientist, who did not want to be identified.

Yet this experience is speedy for China, where reagents, biological materials and other laboratory materials can take weeks or months longer to obtain than in Europe or the United States. With few local suppliers, researchers in China rely on deliveries from overseas. But the shipments regularly get hung up in customs and in a network of distributors, and product quality is sometimes compromised by the tortuous route.

The reagents problem could crimp China's ambitious plans to become one of the top five biotechnology powers by 2020. The Chinese can build a five-star hotel, a highway or a research institute before they can get a mouse from Boston to Beijing. "If a researcher from the United States and I have the same

idea at the same time, I won't even try," says Duanqing Pei, director of the Guangzhou Institute of Biomedicine and Health. "I'd never be able to keep up."

Chinese scientists like to say that a 10-million renminbi (US\$1.5-million) grant in China will go as far as a \$10-million grant in the United States, because labour and overheads are cheaper. But reagents remain some of the biggest expenses — and hassles.

This year's Beijing Olympics made a tough situation even worse. Because of tightened security restrictions, the China branches of Sigma Aldrich, one of the world's biggest suppliers of laboratory reagents, stopped importing from US plants between early August and mid-September. As a result, its China sector was the only one that failed to meet target goals for the company for the third quarter of 2008. Employees at Stealth Peptides International had to drive for two days to bring peptides from its Shanghai headquarters to a contract laboratory in Shenzhen, says chief executive Dajun Yang. In Beijing, neuroscientist Minmin Luo of the National Institute of Biological Sciences couldn't get oxygen in compressed tanks,

or the radioactive probes to do Southern blots. "It was very bizarre," he says.

Now things are mostly back to normal — that is, normal for China. Researchers share stories of waiting a year for a shipment of basic materials. Yuqiang Ding of the Institute of Neuroscience in Shanghai says that part of the problem is that distributors pile up orders and send them all together to reduce shipment fees. Animals, cell lines or reagents with bovine serum are particularly slow to arrive because of strict regulations and confusion at customs. Knockout mice take six months to arrive from most places, says Luo. Shipping from the United States is often a little faster — about five months — because the greater number of orders has made the paperwork more routine.

The delays mean that interesting leads turned up during experiments frequently have to wait. "Sometimes you just can't plan ahead," says Luo. He often ends up saving the exciting parts of the experiment for later and doing controls — experiments that might not even be needed if the more critical parts don't turn out — until he can get what he needs.



GOT A NEWS TIP?

Send any article ideas for Nature's News section to newstips@nature.com

K. CAMPBELL/GETTY

Some researchers have to make substitutions using less-than-ideal alternatives. Luo has applied several times for the right to obtain ketamine from Beijing manufacturers, to no avail. The drug, which can be used recreationally, is tightly controlled in China. Instead he uses phenobarbiturate, which is more difficult to dose than ketamine. "We end up killing some mice," he says. "The experiments go much slower."

Multinational pharmaceutical companies and contract research organizations, which have poured hundreds of millions of dollars into research units in China over the past few years (see *Nature* 455, 1168–1170; 2008), are not immune to the shortages. Sometimes they innovate around the problem. In one instance, Shanghai-based researchers at GlaxoSmithKline took just a week to express and purify more than 10 milligrams of a protein for assay. It would have cost more than \$200,000 and taken at least a month if ordered from a US vendor, says Jingwu Zang, head of the research centre.

But that only works for some reagents. GlaxoSmithKline wants to use its influence and size to sort out the rest. The company intends to increase its staff in China from 200 to 1,000 over the next decade and is pushing local vendors to stock reagents nearby, and others to expedite shipping. It has paid off: some shipments are down to seven days from the United States and one day within China. Zang says he also plans to collaborate with other pharmaceutical companies, many of which are now represented in Shanghai, to establish a centralized local stock facility for all users.

In Guangzhou, Anlong Xu, the vice-president for research at Sun Yat-Sen University, is working to build expertise at the custom houses and exchange agencies that must test and approve imported reagents. "They often don't know how to test the materials, and so they leave them overnight", by which time the dry ice has gone, Xu says. "We can help them." But pulling together the regulatory authorities is a headache; Xu has dedicated one assistant entirely to the task. Luo suggests a quicker alternative: a fast-track system that would accelerate shipments for accredited research institutions.

None of these innovations can arrive too soon for researchers. "Sometimes I feel so frustrated I wonder if I can still do research here," says Luo, who spent nine years in the United States. "It just goes so much faster there."

David Cyranoski

European funding plan 'unviable'

The credibility of the European Commission's initiative to promote research collaborations between academia and industry is under threat because universities cannot afford to take part.

The Commission is setting up six public-private partnerships — known as Joint Technology Initiatives, or JTIs — that will undertake applied research under the European Union's €50-billion (US\$63 billion) Seventh Framework Programme (FP7), which runs until 2013. Roughly half of the money for the JTIs comes from European Union funds, with €3.17 billion paid by member states as part of their contribution to FP7. Industry matches the contribution in kind by offering researchers, lab space and equipment.

But some of Europe's top universities say the terms of the JTIs make it too costly for them to take part. One of the main concerns is that universities can claim only 20% of the indirect costs of these research projects — including infrastructure, computing support and administrative services — from any of the JTI grants. By contrast, other FP7 projects reimburse universities with up to 60% of their indirect costs.

The University of Oxford, UK, has expressed an interest in five projects in the Innovative Medicines Initiative (IMI), which aims to speed up the drug-development process and will receive €1 billion in public funds between now and 2017. But Linda Polik, European adviser at Oxford's research services unit, says the university may have to pull out because it cannot afford to take part. "Even

before the economic downturn we were not in a position to subsidize research to this extent. It is not financially viable to participate in the IMI."

The university complained to the Commission in September, but Polik says that the reply — from European research commissioner Janez Potočnik, who has personally backed the JTI scheme — failed to address their concerns.

University College London has expressed interest in only one IMI research project because the conditions are so unfavourable, says Michael Browne, the university's head of European research and development. "I am not going to be shouting from the rooftops to get our researchers involved in these," he says, adding that the schemes "will definitely fail without university participation".

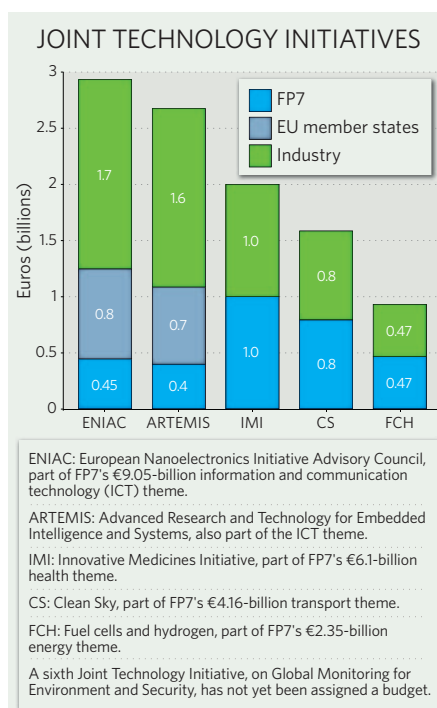
Herbert Mütter, vice-dean for research at the University of Tübingen in Germany, adds that the JTIs have failed to overcome the difficulties of getting universities to collaborate with industry, such as the different styles of conducting and communicating research, and disagreements over intellectual property rights.

The European Association of Research and Technology Organisations (EARTO), which represents public and private research bodies, warns that universities and smaller businesses could "turn their backs on JTIs — as many of them did during the first IMI call for proposals — if the conditions offered are not realistic".

Many research-policy experts contacted by *Nature* expressed the view that the JTIs are being used to subsidize industry, giving them substantial control over the intellectual property generated by the research. The Commission and industry partners involved in the IMI have agreed to set up a working group early next year to discuss the universities' concerns about these rules.

A spokesman for the European Federation of Pharmaceutical Industries and Associations in Brussels — which has five member companies on the IMI's governing board — says it is pleased with how the IMI is progressing, as it proves that industry and academic institutions can work together. The Commission says the decision to place a 20% cap on overhead costs is "a clear decision to focus as much of the funds as possible on research".

"The idea is to get research results to be applied quickly, which requires different rules to those adopted in the rest of FP7," a spokesman says. He adds that the Commission was "satisfied" with levels of interest in the IMI's first call for proposals, which saw 138 submissions. "That strikes me as quite underwhelming," says Christopher Hull, secretary general of EARTO. Natasha Gilbert



SOURCE: EUROPEAN COMMISSION

Space agency funding defies downturn

THE HAGUE, THE NETHERLANDS

"Just look at the smile on my face," beams David Southwood, head of science and robotic exploration at the European Space Agency (ESA), minutes after the outcome of a crucial summit between the ministers of the agency's 18 member states.

Despite the financial crisis, the ministers agreed to commit €10 billion (US\$12.8 billion) to ESA, a 21% increase on the sum promised at the last ministerial meeting in 2005. The funding secures almost everything that was on ESA's wish list, including a spending increase for its flagship science programme.

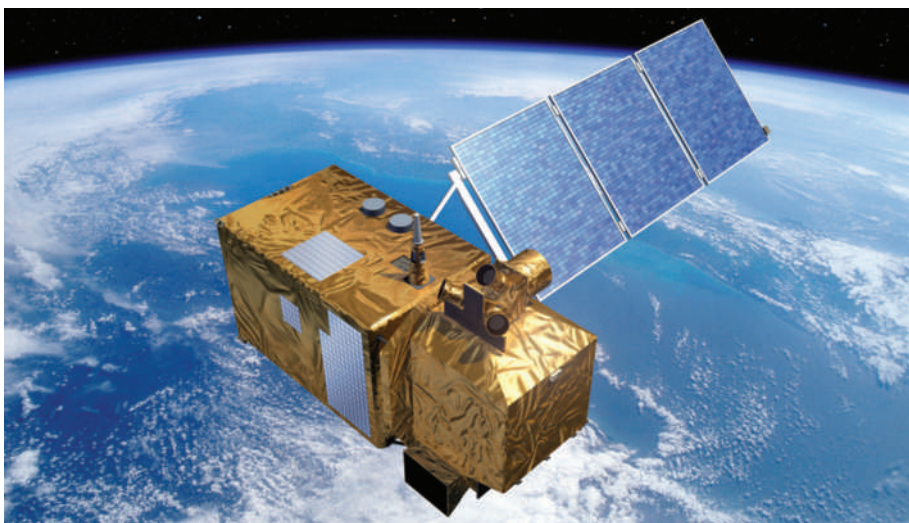
The summit allocated €831 million (to 2018) for a joint ESA–European Union (EU) programme dubbed the GMES/Kopernikus, a network of satellites that would make Europe the world leader in Earth monitoring. It also allocated €943 million (to 2020) for next-generation meteorological satellites; €1.37 billion to operate and carry out experiments on the International Space Station; and €357 million (to 2011) to start design studies for a series of more powerful European space launch vehicles. The meeting was a "complete success", according to a visibly relieved Jean-Jacques Dordain, director general of ESA.

No final decision was reached on ExoMars, an ambitious €1.2-billion plan to send a rover to Mars. But cancellation — which some had feared possible — was avoided. Ministers agreed to ESA's proposed postponement of the launch from 2013 to 2016, but gave the programme €850 million, enough for work to begin, while capping ESA's overall contribution at €1 billion up to its launch. A final funding decision will be taken in September 2009 following a technical review and discussions of how much NASA and other international partners can contribute to make up the shortfall in the programme's budget.

Space for citizens

ESA ministerial meetings are the high point of Europe's space calendar. The multiannual budgets that they fix give projects long-term stability, but can also mean make-or-break for missions. They can be nail-biting affairs for those with missions on the table, often ending in late-night negotiations to hammer out deals.

This year's meeting, held in The Hague in the Netherlands on 25 and 26 November, was upbeat from the start, however. The summit backed 30 new programmes, twice as many as at the last ministerial meeting in Berlin. And ESA will open its first research centre in Britain,



P. CARRIL/ESA

The planned GMES Earth-monitoring programme has won funding from the European Space Agency.

with a facility on the Harwell innovation campus near Oxford focusing on climate-change science and space robotics.

The current financial crisis helped win consensus at the summit, ministers say. "We must invest in the future," says Valérie Pécresse, France's science and education minister. "It is one of the answers to the crisis."

There had also been strong political impetus to make the meeting a success following an EU endorsement this September that space was an important priority. But political support has its price. Pécresse made it clear that governments want ESA's emphasis to be on space activities that benefit citizens, such as the Galileo navigation-satellite system and the GMES/Kopernikus programme.

GMES/Kopernikus will use several families of satellites to provide real-time data on, for example, air quality, water pollution and natural disasters, while generating cross-calibrated, long-term data sets on the state of the planet and its atmosphere.

The next phase of the project, the launch of the second wave of satellites, which was initially budgeted at €1 billion until 2017, was threatened a few weeks ago when Italy reduced its 25% stake to 14% (€120 million). Italy and France also argued that the costs of the secondary 'b-unit' satellites in each family should be borne by the EU (see *Nature* 455, 1013; 2008).

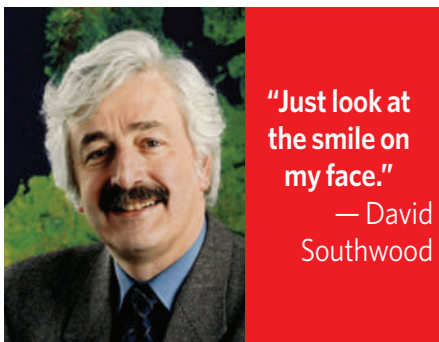
ESA compromised by asking for €857 million, with Britain increasing its share from 1.5% to 10% (€87.5 million). The b-units will still be built by ESA, with the EU paying for their launch — delayed by a year to start from 2015 — and in-orbit validation.

ESA also won its requested 3.5% annual boost to its science-programme budget, slightly more than it negotiated in 2005. Southwood is satisfied with ministers' consistent support, which was so often lacking in the 1990s when space science was a frequent target for cuts.

This keeps all of ESA's planned science missions on track, including the Herschel/Planck space telescope and the BepiColombo probe to Mercury. It will also allow two more missions to launch around 2016, says Southwood. These will be selected in 2011 from four competitors: PLATO, a telescope to search for Earth-like planets; Cross-Scale, to study fundamental space plasma processes; the Euclid dark-matter mapper; and a solar orbiter. "With this money," smiles Southwood, "I can face the world and say, 'Look, we aren't kidding in space science.'"

Declan Butler

See also Editorial, page 546.



"Just look at the smile on my face."

— David Southwood

Melanoma in mice casts doubt on scarcity of cancer stem cells

The rare 'cancer stem cells' thought to seed cancer growth are not so rare after all, according to researchers in Michigan. If they're right, then the strategies used for some cancer therapies may need to change.

The research challenges the notion that there is an elite set of tumour cells that continually regenerates a cancer, says William Kaelin, who studies tumour-suppressor proteins at the Dana-Farber Cancer Institute in Boston, Massachusetts, and was not affiliated with the study. "If it turns out that the cells were never rare, then we've been studying those cells all along," he says. "And a cancer stem cell is just a cancer cell."

Previous studies had transplanted human cancer cells into immunocompromised mice that still had enough of an immune system to fight the transplanted cells. The latest work introduced human melanoma cells into a more severely immunocompromised mouse strain^{1,2}. Under these conditions, the rate of tumour initiation surged. In less immunocompromised mice, only 1 in 837,000 cells regenerated melanoma. In the new mice 1 in 4 cells did

Sean Morrison, who led the work at the University of Michigan in Ann Arbor, warns researchers to interpret his results with caution. "There

are some people out there who look at the paper and say that this completely invalidates the cancer stem-cell field," he says. "That's an over-interpretation." He argues that the evidence for cancer stem cells in certain cancers, such as leukaemia, is particularly convincing.

But researchers who report cancer stem cells in solid tumours should be wary of their results, he says, and should repeat their experiments using techniques that could help transplanted human cells survive in mice. Otherwise, it's unclear whether they have detected a rare subset of cells that can propagate tumours — or simply a rare subset of human cells that can establish themselves in mice.

Cancer stem cells were first identified in leukaemia³ in 1994, by John Dick of the University of Toronto, Canada. Since then, they've been reported in various cancers including those of the brain, breast, colon and pancreas. Researchers agree that techniques used to identify cancer stem cells — sorting cells from a human cancer and transplanting them into mice — are not ideal. Cells that might initiate tumours in patients might not show up in mouse-based assays.

In fact, the rarity of cancer stem cells was called into question last year⁴ by a team of researchers led by Andreas Strasser at the Walter and Eliza Hall Institute of Medical Research in Melbourne, Australia. In mice engineered to overexpress an oncogene, the team showed that leukaemia cells readily established cancers when transplanted into other mice. But other scientists questioned whether those results would apply to spontaneously occurring cancers, human cancers and solid tumours.

Some proponents of the cancer stem-cell hypothesis say they wonder whether melanoma, as studied by Morrison, is unusual in that it might have a particularly high frequency of tumour-initiating cells or an increased likelihood of being cleared out by the immune system. Most importantly, these researchers say that self-renewing cells don't need to be rare to be worthy of special attention. After all, most cancer tumours contain a mixture of cells, and it is still possible that a subset of tumour cells could generate the other cells. "We need to figure out the different functions of the cells in the cancer," says Peter Dirks of the Hospital for Sick Children in Toronto, who has identified cancer stem cells in the brain.

The work has implications for targeting cancer therapies — in particular, whether to develop drugs that go after cancer cells' ability to proliferate rapidly, or to attack the cells seeding the tumours. Michael Clarke, a stem-cell biologist at Stanford University in Palo Alto, California, says that current therapies work to block proliferation. Some experimental therapies that block stem-cell pathways are in clinical trials, such as those being developed by Genentech in South San Francisco, California, and Merck of Whitehouse Station, New Jersey. Clarke and others think it is essential to identify and purify cancer stem cells to work out the best pathways to target.

Other researchers, such as Strasser, see little use in chasing special subsets. In those cases where the cancer doesn't return, he says, "it's where we are very good at killing all the cells of the tumour". ■

Monya Baker

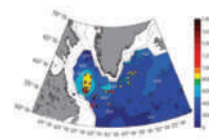
1. Quintana, E. *et al. Nature* **456**, 593–598 (2008).

2. Eaves, C. J. *Nature* **456**, 581–582 (2008).

3. Lapidot, T. *et al. Nature* **367**, 645–648 (1994).

4. Kelly, P. N., Dakic, A., Adams, J. M., Nutt, S. L. & Strasser, A. *Science* **317**, 337 (2007).

See also Authors, page xiii.

**OCEAN MIXING**

North Atlantic overturning restarts after stagnant decade.

www.nature.com/news

Can triniobium tin shrink accelerators?

The superconducting cavities that drive most of the world's particle accelerators are running out of room for improvement. But new theoretical work suggests that overlooked superconducting materials could be used to make cavities that accelerate particles to higher energies over shorter distances — thereby doing the job more cheaply.

Still, it would take years to develop and test new accelerator components made of materials such as triniobium tin.

For decades, researchers have worked to improve the performance of superconducting cavities. When stacked in a row and pumped with microwave pulses, they build up large electromagnetic fields that accelerate the charged particles used in particle-physics experiments, or in synchrotrons working as high-intensity light sources. The cavities are typically made of niobium, a metal that superconducts when cooled to a few degrees above absolute zero.

Slow and steady

Scientists have also grappled to improve the performance of the metal, working with different welds, rolls, cavity shapes and surface treatments to make steady but slow gains to the accelerating gradients. The gradients, measured in megaelectronvolts (MeV) per metre, are a measure of how quickly cavities can push particles up to a particular speed. The best laboratory performance of a single cavity is 59 MeV per metre; full machines rely on strings of hundreds of the cavities and would be hard-pressed, under real-world conditions, to achieve gradients half as good.

Other superconducting metals, such as magnesium diboride and triniobium tin, have been discovered, but there have been few efforts to make cavities with them. "People have been assuming that these new materials would be much worse, or not much better than niobium," says James Sethna, a condensed-matter theorist at Cornell University in Ithaca, New York. "Our estimates suggest that you could do dramatically better."

Sethna's work, done with Gianluigi Catelani of Rutgers University in Piscataway, New



Particle-accelerator cavities are currently made from niobium.

Jersey, will appear in *Physical Review B* and is available on the physics preprint server arXiv (G. Catelani and J. P. Sethna <http://arXiv.org/abs/0810.4720>; 2008). Sethna says that the peak gradient for triniobium tin is 120 MeV per metre and magnesium diboride could reach 200 MeV per metre.

Although a distant goal, achieving such gradients could result in huge savings for future accelerators. For example, the US\$7-billion International Linear Collider (ILC) — a wished for next-generation particle accelerator — will

need thousands of cavities, stretching along a tunnel 31 kilometres long, to help it produce energies of 500 gigaelectronvolts. And that's if its niobium cavities can reach ambitious target gradients of a bit more than 30 MeV per metre (see graph). Ramping up the gradients to a theoretically possible peak of 200 MeV per metre could significantly reduce the length of the ILC, therefore also reducing costs of most of its physical parts such as tunnels and beamlines.

Breaking the limits

According to Sethna, old theories describing the limits of superconducting materials didn't correctly account for ultracold operating temperatures or vortices in the superconductors, which, in creating barriers to the magnetic field lines, allow the cavities to operate at higher gradients. "The excuse has been, 'We've worked so hard on niobium, why try these other materials?'" says Sethna. "All of a sudden, we're telling them, 'you've been listening to the wrong theorists.'"

Hasan Padamsee, a superconductivity expert at Cornell University, is excited by Sethna's work. In the late 1990s he made a test cavity out of triniobium tin, and it performed less well than niobium alone. But the new work gives Padamsee confidence that the result was due to a problem with the particular sample rather than a fundamental quality of the material.

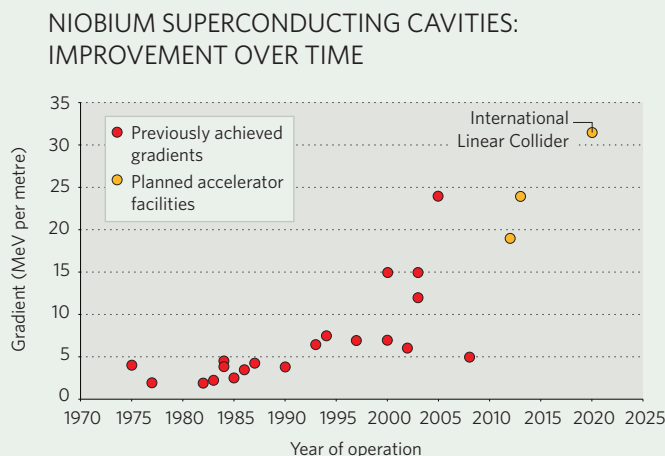
He says a good working cavity with the new materials would take at least five years to perfect. That's too long to help the ILC in the near term,

but perhaps soon enough to be incorporated into upgrades to it.

Peter Kneisel, a senior staff scientist at the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, is less optimistic. He would prefer to wring incremental advances out of niobium cavities, and accept their limitations.

"The grass on the other side of the fence is always greener," says Kneisel, who has spent 40 years working to improve niobium. "It will take, I don't know, 40 years again to bring it to some practical use in accelerators."

Eric Hand



Europe to pay royalties for cancer gene

BRCA1 patent decision may be ignored in clinics.

It has been one of the toughest, and most impenetrable, biological patent cases in the history of the European Patent Office. But the University of Utah in Salt Lake City has finally won its battle to keep some European patents on *BRCA1* — a gene associated with breast and ovarian cancer.

The ruling means that the patent owners now have the right to collect royalties on tests carried out on tens of thousands of women across Europe every year. The tests identify those who have a high risk of developing breast cancer because they have a mutated *BRCA1*.

"It is disappointing after our 7-year fight even though we have managed to reduce the scope of the patents," says Dominique Stoppa-Lyonnet, a clinical geneticist at the Curie Institute in Paris. "We had all been freely testing the *BRCA* genes since they were first described in the early 1990s," she says.

Between 10% and 15% of all heritable breast and ovarian cancers have a mutation in *BRCA1* or *BRCA2*, another gene linked to breast cancer, leading to as many as 5,000 new cases of both cancers in the European Union every year.

An inexact sequence

Clinical geneticists do not agree with monopolies on diagnostic testing of genes for such diseases because they believe they block the competition that could lead to the development of better, cheaper products. Some of those contacted by *Nature* say that they will continue to test for the mutations in defiance of the patent.

Gert Matthijs, who heads the Centre for Human Genetics at the University of Leuven in Belgium, says that no European clinic has paid royalties for *BRCA1*-related diagnoses since the European Patent Office first granted three *BRCA1* patents to the biotech company Myriad Genetics in 2001. Myriad, which is based in Salt Lake City, transferred ownership of the patents to the University of Utah in November 2004. It was facing growing opposition to its patents from genetics societies, research institutes, hospitals and some governments.

In 2004, the European Patent Office revoked the key patent covering any mutations in *BRCA1*, as well as any diagnostic method for their detection. Its oppositions board, which considers

patent challenges, was convinced by arguments that Myriad had registered an incorrect gene sequence when it first filed the patent in 1994 — and that a perfect reference sequence is required to make a full diagnosis.

But the office's highest board of appeals reversed this decision on 19 November, after the patent owners said they would reduce the scope of the patent to cover only 'frame-shift mutations'. These mutations — usually the deletion or insertion of one or two nucleotides — lead to the wrong sequence of amino acids being translated from the gene.

This board, whose ruling is final, agreed that an exact sequence is not required to detect these 'frame-shift' mutations — which account for around 60% of those associated with breast and ovarian cancer.

But clinical geneticists say that the ruling is confusing because the patent now applies only to some of the genes picked up during diagnosis. "It becomes unclear at which point during the analysis practitioners like myself would start to infringe the patent," says Matthijs. The

University of Utah declined to say whether it will now seek out and challenge infringers, which it had not done while the patent oppositions were ongoing.

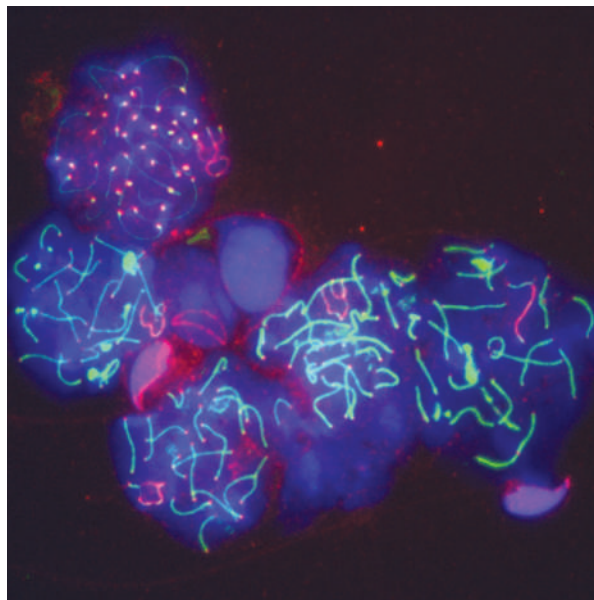
The patent owners have chosen to pay the fees needed to maintain the patent only in large European countries. It will not apply in Belgium, for instance, says Matthijs, "but it is applicable in neighbouring France and the Netherlands, where testing could now start to cost a lot more".

Myriad holds very broad patents on both *BRCA1* and *BRCA2* in the United States, where it charges US\$3,120 for a full analysis of both genes and \$460 for a single mutation test. The single test is used mainly by members of families in which a mutation has been detected before. In Europe, the tests are usually done in large academic institutions and hospitals at a cost of up to €1,500 (\$1,900) for both genes. "We will wait to see what royalties the University of Utah might demand of us, but it won't stop us testing the gene in France," says Stoppa-Lyonnet.



"The patent ruling won't stop us testing the gene in France."

— Dominique Stoppa-Lyonnet



Mutations in the gene coding for *Brca1* (red) raise the risk of developing breast cancer — for both mice (above) and humans.

Patent patchwork

William Hockett, a spokesman for Myriad, says that the company is keen to discuss the test's costs with national health systems, pointing out that "our analysis is covered by insurance in the United States because it is highly cost-effective to the insurers".

The patent office granted Myriad a broad patent on *BRCA2* in 2003. That patent was reduced in scope to cover just one mutation, common in Ashkenazi Jews, in 2005. Myriad is challenging a second, broader, patent on *BRCA2* that is co-owned by the charity Cancer Research UK. Mike Stratton of the Wellcome Trust Sanger Institute in Cambridge, UK, who is one of the inventors behind the patent, says that they "took out the patent to defend the gene against other patent approaches. We offer free licensing to any reputable laboratory who wants to use it," he says.

Alison Abbott

S. LAURE/INSTITUT CURIE

MRC NIMR/WELLCOME IMAGES



HAVE YOUR SAY
Comment on any of our
news stories, online.
www.nature.com/news

Austrian ethics watchdog launched

In the wake of the misconduct scandals that have shaken the scientific community in Austria, universities and leading research institutions have created an Agency for Scientific Integrity.

In future, a scientist at any Austrian university or research institute who is concerned about misconduct will be able to turn to this body, the first of its kind in the country, for confidential help. Six scientists from outside Austria will sit on its board, to prevent any influence from the strong local networks that exist in the small country, along with just one Austrian legal adviser. The board members, who will be named in the next few weeks, will determine the organization and procedures of the final agency, which should be operational from early next year.

Christoph Kratky, president of Austria's basic-science funding agency the Austrian Science

Fund (FWF), spearheaded the effort to establish the agency. Plans for such a body had been gestating for a couple of years, he says, but two recent cases have accelerated its birth.

One case involved research from the Medical University of Vienna purporting to show that electromagnetic fields from mobile telephones caused DNA damage. An internal university

investigation concluded in August that data in two papers had been fabricated.

That affair came hot on the heels of an even bigger scandal at the Medical University of Innsbruck, which involved

clinical trials of an experimental procedure for stem-cell therapy to treat urinary incontinence, led by urologist Hannes Strasser. The government Agency for Health and Food Safety later described the trials as illegal and invalid because the researchers had failed to get appropriate approval (*Nature* 454, 922–923; 2008),

and on 6 September *The Lancet* withdrew Strasser's 2007 paper, which had reported clinical success of the therapy (*Nature* 455, 150; 2008). The university failed to investigate the affair, and an enquiry by the Austrian National Academy of Sciences is ongoing, although such activity is not within the academy's remit.

The new agency will be useful in cases in which local connections serve to cover up misbehaviour, says Kratky, who is also a structural biologist at the University of Graz. He originally started pushing for an agency because the FWF gets occasional tips about plagiarism from grant reviewers. "We typically convey them to the university involved, but only some investigate properly," he says. At first the universities were sceptical, he says. "But after the Innsbruck and Vienna cases blew up, the research minister Johannes Hahn said that he would create an agency himself if they didn't do it."

Alison Abbott

"The research minister said he would create an agency himself if they didn't do it."

UK charity launches assault on deadliest cancers

Cancer Research UK is to fund more work on pancreatic, lung and oesophageal cancer and will open up to 20 more research centres as part of its plans to disburse £1.5 billion (US\$2.3 billion) over the next five years. The charity, the world's largest independent

funder of cancer research, already provides core funding for seven dedicated cancer institutes in Britain.

Compared with other cancers, such as prostate and breast cancer, for which survival rates have improved, "essentially no progress" has been made in preventing deaths from pancreatic, lung and oesophageal cancers, says Harpal Kumar, Cancer Research UK's chief executive.

The charity recently delayed making decisions on applications for project grants, due in January 2009, until April. But it will press ahead with its five-year strategy commitment despite the financial downturn. "We are weathering the storm," says Kumar.

For a longer version of this story, see <http://tinyurl.com/5lulga>

Greenhouse gases hit modern-day highs

Atmospheric concentrations of greenhouse gases reached new highs in 2007, according to the most recent analysis by the World Meteorological Organization (WMO).

Concentrations of carbon dioxide, methane and nitrous oxide — which together contribute 88% of the anthropogenic global-warming effect — were last year 37%, 156% and 19% above pre-industrial levels, respectively.

Since 1990, total radiative forcing — the re-radiation of heat back towards Earth's surface — by all long-lived heat-trapping gases has increased by 24%, the WMO reports. In 2007, the average concentration of carbon dioxide rose to 383 parts per million, an increase of 0.5% from 2006. Methane and nitrous oxide concentrations also rose, reaching new highs of 1,789

Europe's synchrotron sharpens up its X-rays

The European Synchrotron Radiation Facility (ESRF) has won approval to begin a €177-million (US\$224-million) upgrade of its facility in Grenoble, France (pictured). On 25 November, the ESRF council, which represents 19 member states, approved funding for construction to begin on eight new beamlines for the facility. The new lines will deliver sharply focused beams of X-rays just tens of nanometres in diameter.

Among other applications, researchers plan to use them to image the nanometre-scale features of various materials and follow the progress of chemical reactions. Other synchrotrons have achieved similar-sized beams, but the ESRF's will be the first available for general use by scientists. Construction of all eight lines is expected to run until the end of 2015, but the first ones could be ready as soon as mid-2013.



D. MOREL/ESRF

M. RIOPA/AFP/GETTY IMAGES

and 321 parts per billion. The increases are in line with the Intergovernmental Panel on Climate Change's predictions for greenhouse-gas concentrations in a world experiencing rapid economic growth.

Europe rejects Wisconsin's key stem-cell patent

Human embryonic stem cells whose derivation requires the destruction of embryos cannot be patented, the appeals board of the European Patent Office (EPO) ruled on 27 November (see also 'Europe to pay royalties for cancer gene', page 556).

The decision concludes a 13-year battle over a patent application from the Wisconsin Alumni Research Foundation for a method to derive stem cells from the undifferentiated cells of human embryos. The patent, which has been granted by the US Patent and Trademark Office, would also have covered some of the most commonly used human embryonic stem-cell lines.

The EPO decided that the patent contravenes a clause in its own convention, which says that anything whose exploitation is considered "contrary to public order or morality" cannot be patented. The ruling will not affect patents or patent applications on adult stem cells.



Stephen Hawking is taking a position in Canada.

Hawking plans long commute to Canada

Theoretical physicist Stephen Hawking, soon to retire as Lucasian professor of mathematics at the University of Cambridge, UK, has accepted a distinguished research chair at Canada's Perimeter Institute for Theoretical Physics in Waterloo, Ontario. The post is a travelling position, which means Hawking can stay on as an emeritus professor at Cambridge while making regular visits to Canada.

The executive director of the institute, Neil Turok, who took up his post in October, is a former Cambridge colleague of Hawking's. Following his retirement from Cambridge at the end of the current

academic year, Hawking, who is almost completely paralysed by amyotrophic lateral sclerosis, will begin his visits to the Canadian institute next summer. Applications for the Lucasian chair will close on 15 December.

Chemicals firm cancels plans for biodiesel plant

The world's third-largest chemicals company, Ineos, has shelved plans to build four large European biodiesel plants because of the "continued and prolonged global economic downturn".

The plants — in Belgium, France, Germany and Britain — were intended to come on stream over the next four years. Together, they would have produced a total of 2 million tonnes of fuel annually.

Ineos is the largest private company in Britain and was founded just over a decade ago. The company's rapid growth has been funded by debt financing, which it used to acquire more than 70 chemical-production sites across 14 countries.

The company says it is now cancelling all non-essential expenditure, and recently announced that it hoped to renegotiate the terms of loans totalling more than £5 billion (US\$7.4 billion).

The Sputnik fable

Oversimplifying the effect of the space race on US science funding could lead scientists down the wrong path, says **David Goldston**.

In these days of economic uncertainty and political transition, science advocates understandably seek guidance (and solace) from an earlier age when, supposedly, the issues were clearer and success was assured. As a result, a common refrain in speeches and articles is that what's needed to jump-start US research spending now is a 'Sputnik moment'. But this longing for the hysteria that attended the 1957 launch of the Soviet satellite fosters undue pessimism about today — and is as likely to hamper advocates' goals as to spur them.

The problem with the Sputnik trope is that it makes the 1950s look too simple and later periods look too bleak. The false simplicity leads scientists to believe that it was once possible to push for funding without getting tangled in politics, and the false bleakness leads them to conclude that the only way to win public support for science is to posit a crisis and claim to be able to resolve it.

The standard version of the Sputnik fable is a kind of morality tale that goes more or less like this: in the 1950s, the United States was a laggard in science and technology, which was demonstrated starkly when the Soviet Union launched the world's first satellite. In response, the American public and politicians saw the error of their ways and reacted forcefully, increasing investments in research and education across the board. That put the United States in the scientific vanguard for at least a decade until the nation began to forget the lessons of Sputnik, allowing new rivals such as Japan to begin to eat away at its lead.

There is some truth to this narrative; science spending did increase in response to Sputnik, and new agencies such as NASA were created. But all the particulars are wrong, and the story obscures as much as it illuminates. The whole debate around responding to Sputnik was highly politicized, with Congress trying to showcase concrete actions and President Dwight Eisenhower trying to resist what he saw as an excessive response given that US satellite programmes were not actually behind those of the Soviets. Research spending was already increasing before the Sputnik launch, and the increases that followed it were highly concentrated in the areas of defence and space, and allocated more to development than to research.



PARTY OF ONE

And science and education budgets did not lose momentum as the Sputnik era receded. Later increases in basic research and education budgets were at least as significant as those in the 1950s, but they did not emerge in the same kind of crisis atmosphere. Rather, they resulted from the overall growth of domestic spending that took off with President Lyndon Johnson's Great Society initiatives, and from a variety of new political concerns, including economic competitiveness, the quality of health care and equal access to education. The National Science Foundation, for example, experienced major expansion beginning in the 1970s, and a boom time for the National Institutes of Health began in the 1990s. But despite these increasing investments, the United States faced new economic rivals, including western European nations that had taken years to recover from the devastation of the Second World War.

Keeping it real

If scientists and advocates had a more realistic understanding of Sputnik and its aftermath, they might not despair that a golden age has turned to dross every time science funding becomes a political debate, or when development is favoured over research, or when the growth of science funding turns out to be largely a function of overall federal spending, not an independent variable. They also might be more wary about framing every plea for funds as a claim that the nation is approaching a moment of crisis that science alone can address. That is not the only line of argument that has yielded money for science, and it is an approach that can backfire if promises fail to materialize or the crisis passes. Those waving

around the 2005 National Academy of Sciences report *Rising Above the Gathering Storm*, for example, ought to be careful not to oversell their wares.

It's interesting in that regard that the Sputnik myth tends to ignore the later history of NASA, the satellite's offspring. The space race was the only widely accepted rationale for the agency, and its human spaceflight programmes have been fumbling around for purpose, money and support ever since the cold war crisis began to wane. NASA proponents now wonder why the agency isn't included in plans to increase research spending to improve economic competitiveness, forgetting that the Sputnik moment was tied specifically to the Soviet threat and did not benefit all agencies equally.

The reaction to the Arab oil embargoes of the 1970s — an acute threat that could be addressed in part with technology, arguably the closest thing to a Sputnik moment since the 1950s — offers a similar cautionary tale. Spending on energy research climbed in the immediate aftermath of the crisis, but then plummeted along with oil prices and the belief in government programmes that came under attack from President Ronald Reagan.

Happily, scientists seem to be drawing the right lessons from the 1970s, at least in the energy realm: don't focus excessively on the immediate crisis, don't promise too many benefits too quickly, and don't assume spending on your pet projects can thrive in a political atmosphere that is hostile to government programmes in general. As President-elect Barack Obama is putting together his energy plan, groups such as the American Physical Society (APS) are rightly calling for a mixture of shorter- and longer-term research while acknowledging that the research needs to be coupled with new energy policies if it is to make a difference. (I served as a co-chair for the recent APS study.)

As George Orwell put it, whoever controls the past controls the future. The stories we tell about the past influence how we go about building the future and how we respond to political obstacles today. The Sputnik itself was benign, but the heroic tale that it spawned can cause real damage. Scientists present the story as a heartening and apolitical Aesopian fable: science is first ignored, then the need for it is recognized by all, and then it rides single-handedly to the rescue of the nation. That has left the scientific community far too easily discouraged when it confronts the real political world and needlessly short of stories to deploy on its own behalf.

David Goldston is a project director at the **Bipartisan Policy Center** in Washington DC. Reach him at partyofonecolumn@gmail.com.

FIVE CROP RESEARCHERS WHO COULD CHANGE THE WORLD

The current crisis in worldwide food prices reinforces the need for more productive agriculture. **Emma Marris** meets five ambitious scientists determined to stop the world from going hungry.

The rust hunter

Peter Dodds

Molecular biologist at the Commonwealth Scientific & Industrial Research Organization Plant Industry in Canberra, Australia

Timescale for change: 10 years

It was only when his supervisor showed him a picture of a tractor crossing a rust-infected field of wheat that Peter Dodds really understood what he was up against. Behind the tractor bloomed an orange plume of spores several times higher than the vehicle itself. "It is just amazing the amount of spores that get released in an infected wheat field," says Dodds. "It is like looking at Mount Everest."

The cloud of fungal spores revealed a terrifying strength in numbers. You might think that if a mutation that will overcome wheat's resistance to a strain of rust is a one in a million fluke it would not be worth worrying about. But there in that one picture were thousands of billions of spores, obscuring the sky. And there are millions of fields.

Since the early days of its domestication, wheat has been plagued with various strains of rust. The fungus's spores infiltrate the stomata through which the plant takes in carbon dioxide from the atmosphere, and poke tentacle-like haustoria into the wheat's cells, extracting their nutrients. If the conditions are to its liking, stem rust (*Puccinia graminis*) can kill 50–70% of the wheat in an area.

These losses can be prevented if the wheat is resistant to the rust. Resistant plants can identify an invasion early and sacrifice the invaded cell, stopping the rust in its tracks. But domesticated wheat, which is bred for yield, has a limited gene pool from which to draw the genetic variations that might offer resistance. So whenever a strain of fungus overcomes the resistance genes, researchers need to scour wild relatives of wheat for new ones that can be introduced to the crop by cross breeding.

In 1999, the stem-rust-resistance gene *Sr31* — hitherto an unbeaten champion and relied on by wheat farmers throughout the developing world — succumbed to a

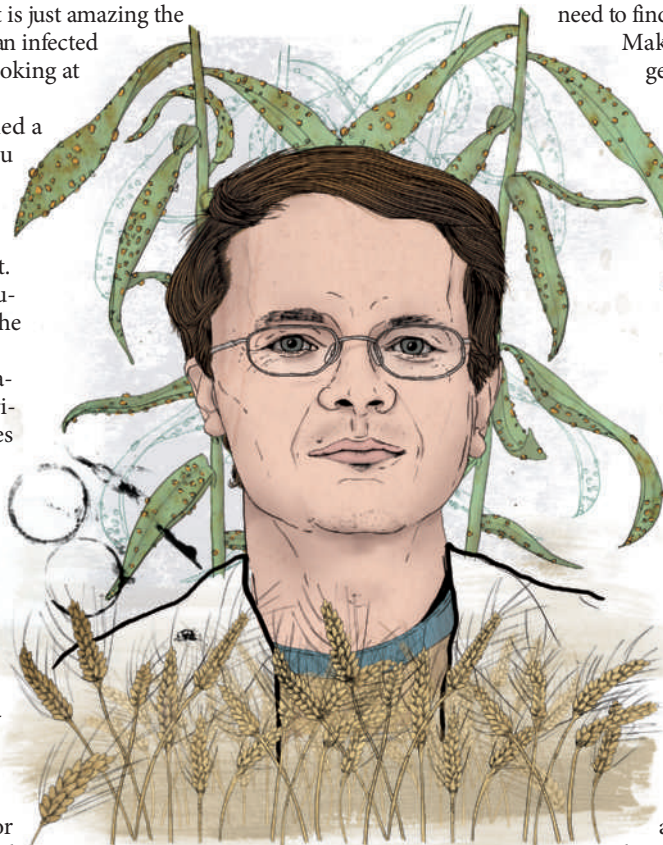
new strain of rust from Uganda. The Ug99 rust has since spread through Kenya and Ethiopia, crossed the Red Sea and reached Iran (see map, overleaf). Experts estimate that 19% of global wheat production grows in the potential migration path of the rust. If it covers the entire zone at risk, losses are estimated to be tens of millions of tonnes, and thus billions of dollars, per year.

Dodds might be the man to stop it in its tracks. A laconic Australian, originally from Melbourne, Dodds wants to understand how rust invades cells so that he can engineer resistance proteins from scratch. That would remove the need to find hardy relatives in their wild redoubts.

Making the leap from crossbreeding to genetic engineering would also remove the problem of incorporating large chromosome segments that reduce yield along with the resistance genes. Wheat with *Sr31*, for example, is unfit for making yeast-based bread thanks to some undesirable genes that come along with its top-of-the-range resistance.

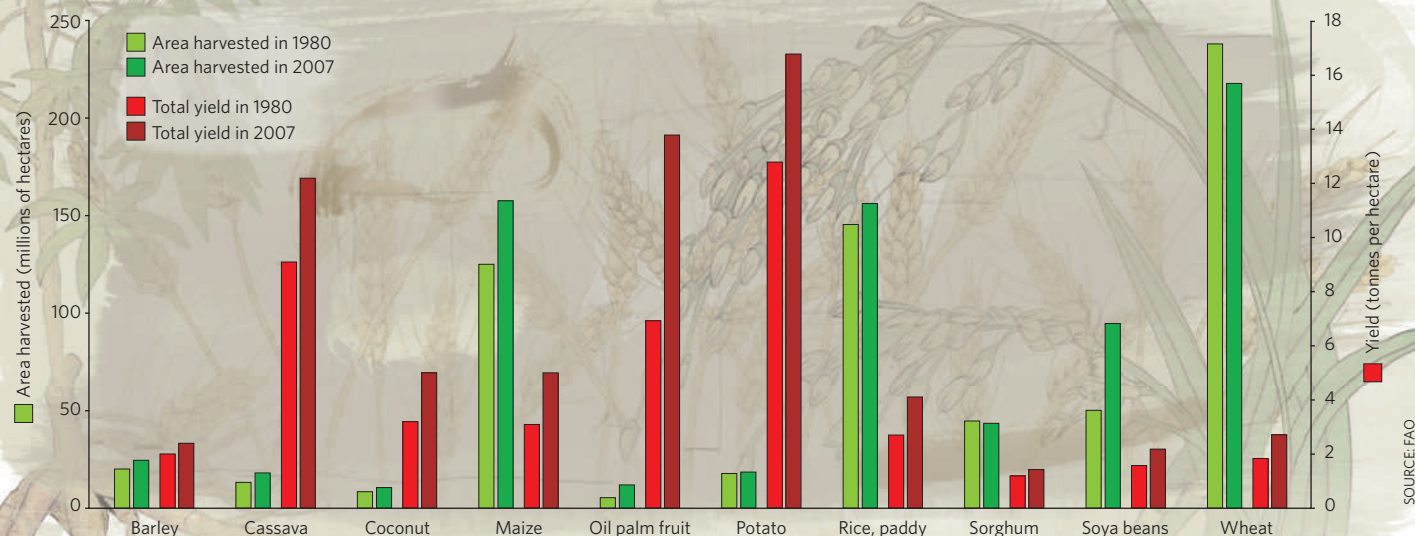
Dodds and his team in Canberra use flax and flax rust as a model; they rarely see a commercial wheat field, and never an infected one — hence Dodds's amazement at the ominous orange wake in that photo. His team's work centres on substances that the rust secretes to gain entrance to plant cells, or perhaps to manipulate their metabolism — substances the plant can use to recognize that is it being invaded, and thus get its countermeasures up and running. These compounds, says Dodds, "make [the rust] vulnerable to recognition by the plant, but the rust can't do without them".

The hope is that once the interaction between these substances and the wheat immune system is worked out, all a bioengineer will need to do is to look at how the rust strain has evolved to identify effective resistance genes. Such an advance would change the arms race between man and rust. "Initially it will make the catch-up several times faster," says Dodds. Eventually, he adds, varieties engineered to express several new resistance genes might be too complex for rust to adapt to for years and years. After that it will just require



ILLUSTRATIONS BY JAY TAYLOR

THE WORLDWIDE HARVEST OF SELECTED FOOD CROPS



occasional tweaking to keep ahead of the pathogen.

But will countries in East Africa, where rust often takes hold, open their arms to genetically modified wheat when it is ready for use? "That is a question that I have difficulty answering," says Dodds. "When any part of the world is going to be ready to accept that is hard to know, but having the things available puts forth a strong argument."

"Dodds took the analysis of the flax-rust interaction to another level," says Jonathan Jones, a specialist in disease resistance at the John Innes Centre in Norwich, UK. Jones says that Dodds is "the best young guy coming through. He's an outstanding young scientist, but also the work that he is doing is of profound importance." Many molecular biologists, says Jones, wouldn't be interested in such a difficult

problem. "The pay off is too long term," he says.

Dodds seems to have the long-term gene. He has been working on rust for 10 years and says he still finds it as interesting as ever. Ug99 has sharpened his focus. "There is certainly a higher level of urgency when you see that there is a problem that needs a solution rapidly. Every now and then you get a really bad epidemic like this and it requires a response."

The perennial optimist

Jerry Glover

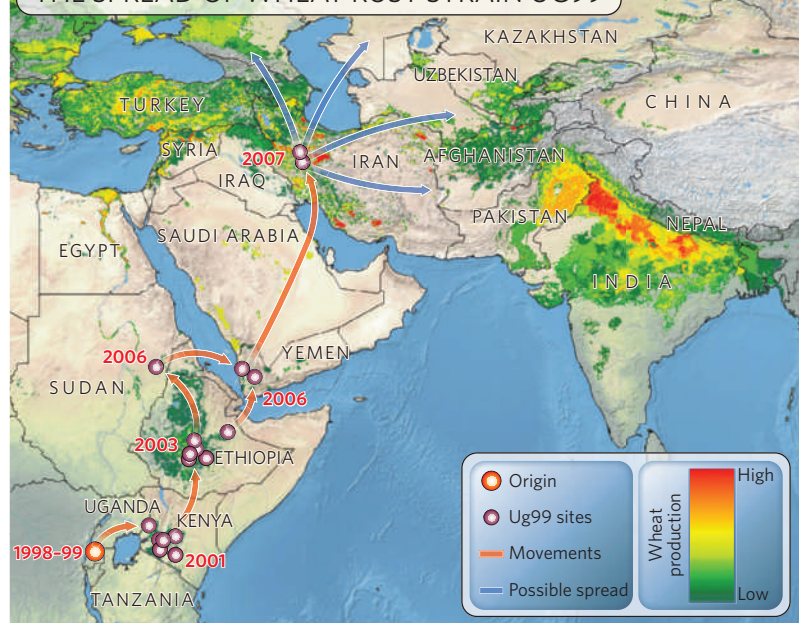
Agroecologist at the Land Institute in Salina, Kansas
Timescale for change: 30 years

He didn't want to stay on the Colorado farm he grew up on; he wanted to become a philosopher. But a summer job in landscaping led Jerry Glover to a community-college course in soil science. He had always loved the loamy smell of the freshly turned land at ploughing time, and the soil drew him back. He got a PhD at Washington State University in Pullman and took a job at the Land Institute, which focuses on sustainable agriculture. There he works on a soil-improvement project that he acknowledges might not be finished in his lifetime.

Glover and his collaborators around the world are in the midst of a decades-long attempt to breed wheat into a perennial plant. About 85% of Earth's cultivated land is planted with annual crops (see world map, page 566). That means that every year they must be planted anew from seed. On most farms that means ploughs, and ploughs mean carbon loss — those rich loamy smells — and erosion. One way round this is 'no-till' farming, in which seeds are inserted into unploughed land. But although this technique has its advantages, when the crops involved are annuals, they never manage to get their roots as deep into the soil as they might if they had longer to grow.

If crops such as wheat could instead be made to persist from year to year — to become perennial — they would require less fertilizer and fewer passes of heavy machinery,

THE SPREAD OF WHEAT RUST STRAIN UG99



and would have more growing seasons. The world's arable lands would revert to something more like the prairies and savannahs that agriculture has replaced, and that would bring many benefits. Some perennial grasslands can be harvested time and again with few or no human inputs of fertilizer without depleting the soil of nutrients¹. With roots like the beards of grand old men, perennials control erosion in ways that the spindly whippersnapper whiskers of annuals simply cannot. They also improve the quality of the soil and pump more organic matter into it.

Glover's particular concern is nitrogen. "We need a lot for good plant yield," he says, "and we often need to remove a lot for our needs." All the nitrogen removed from a farm in the form of protein in crops must somehow be returned to the soil if long-term fertility is to be assured. Most farmers who can afford to do so use synthetic nitrogen fertilizers for the job (see graph, below). But these fertilizers are expensive to make, and are often applied in such a way that a fair amount flows unused into rivers, disrupting the ecosystems downstream.

As the roots of perennials can tap into naturally occurring nitrogen resources that annual crops can't reach, they could make more harvestable protein with less added nitrogen. Also, perennial plants have more opportunities and more root space to interact with the mycorrhizal fungi and bacterial populations that fix nitrogen for them. According to Glover's experiments, perennials, as represented by tall-grass prairie meadows, require just 8% of the energy that a typical high-input annual wheat field needs to make the same amount of harvestable nitrogen.

Glover lives across the road from some of his experimental plots, a few kilometres out of town, with his wife and triplet sons, who are still too young to start running through the test fields after the bugs and birds that find the perennial



plots more congenial than normal fields. Cats and dogs lounge around the property. Comparing his spread to the working farm he grew up on, he says, with a laugh, "I feel more like a gentleman farmer on five acres in the country."

Glover's graduate adviser, John Reganold, says that Glover is motivated by thoughts of the future. "I don't think he would take on a project that would not feed the planet and also be good for the environment," he says. Glover was "the best graduate student I've ever had, and one of the best that I've ever seen," he adds.

"Agriculture is one, if not the largest, single threat to biodiversity in terms of human behaviour," says Glover. "People have to eat — but what can they eat without destroying the environment?" Perhaps they will someday do that by eating food from farms that run more like the natural landscapes they replaced, acting like a healthy ecosystem and a farm all at once. ■

The thriving peasant Zhang Jianhua

Plant physiologist at Hong Kong Baptist University

Timescale for change: now

Unlike most successful agricultural scientists, Zhang Jianhua knows what it is to be hungry. And it was an observation sharpened by hunger that provided the motivation for much of his research.

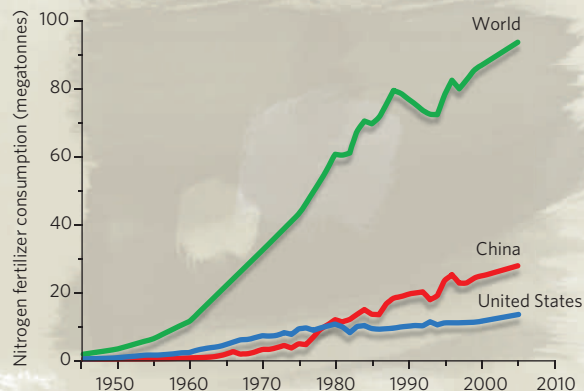
Zhang, 52, grew up on a farm in the Chinese countryside where every day was hard work in the fields, just as it is for the best part of a billion of his compatriots today. His father, a schoolteacher, was imprisoned in a re-education camp for three years starting in 1957, one of tens of thousands branded as 'rightists' after responding to the call for constructive criticism in Chairman Mao Zedong's 'Hundred Flowers Campaign'. Later, Zhang's father was assigned to teach far from his family, and rarely saw them.

Zhang, his siblings and his mother worked on a collective farm, living in a hut with an earth wall and a rice thatched roof. Each household in the collective also had a small plot for its personal use. Zhang vividly remembers the year "when I was supposed to be in high school" when he noticed his family plot, which sat a little higher than the others the irrigation channels served, was drying out a bit just as the rice crop was producing its grains in mid-autumn. "I worried a lot that my rice's yield would suffer," he says. "But at harvesting time, I found that the kernel weight was actually heavier than average. I always remembered that."

Zhang's abilities as a farmer saw him promoted to a post as a technical expert in his collective. He went on to a local agricultural college, where he studied crop production and breeding, and taught himself English in part by reading works by Charles Dickens with a dictionary in his left hand. "I spoke with a very strange accent," he says, laughing.

SOURCE: V. SMIL

THE RISE OF SYNTHETIC NITROGEN



In 1985, the Chinese government offered him an opportunity to work abroad and he wound up in Bill Davies's plant-physiology lab at Lancaster University, UK. "There was something pretty compelling about his letter," remembers Davies. "He was the first Chinese person to join the lab." Worried about his English, Zhang avoided answering the phone, but started pumping out publications. "He did the definitive work² showing that roots can signal to shoots using a particular plant hormone — abscisic acid — in response to drought stress," says Davies.

One idea Zhang worked on, as remembered from his rice patch back at home, was that drought stress can, at times, convince a plant to throw all its resources into reproduction, as death by drought might be imminent. As a result, nutrients throughout the plant are rushed to the grains. Management techniques derived from this insight, which Zhang is now studying in molecular and genetic detail, go by the name of 'deficit irrigation'.

Zhang didn't invent deficit irrigation, but he is now one of its most influential scientific proponents. "As the result of his work," says Davies, "it became a much more focused effort." These days, Zhang's team is looking at the gene functions that are induced by water stress. The team has found, for example, that when deprived of water, plants protect themselves by producing more enzymes that scavenge reactive oxygen species. The group is now busy untangling the details of that regulation.

On the management side, one of Zhang's specialties is a technique called 'partial root zone drying', in which some roots are watered and others are not. The idea is that the plant gets both the water it needs and the focus-your-efforts, put-it-all-into-seeds-for-tomorrow-we-die hormone signals. This

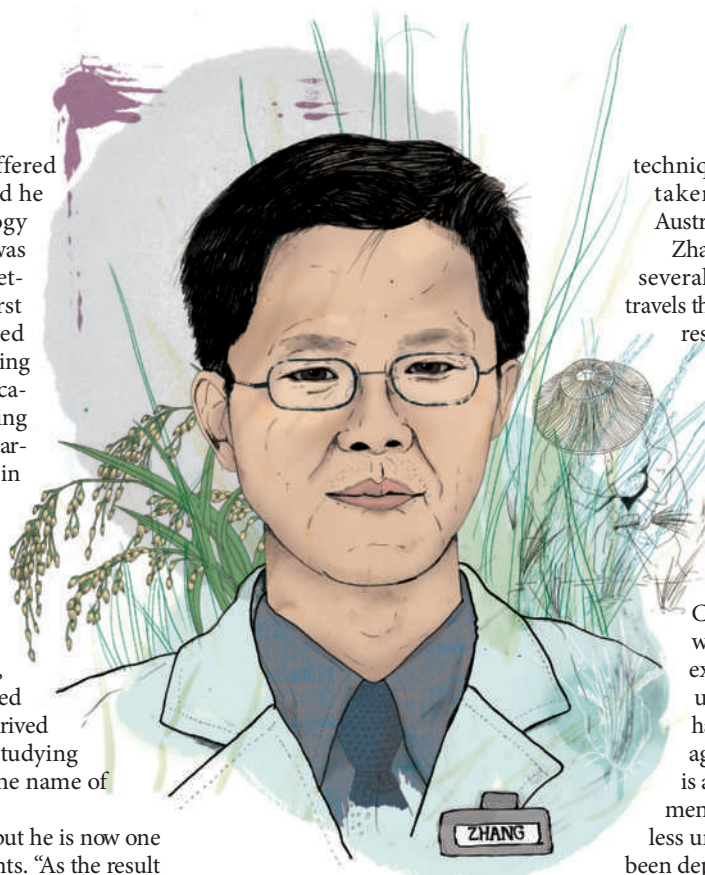
technique has been enthusiastically taken up by wine growers in Australia, among others.

Zhang moved to Hong Kong after several years at Lancaster and now travels throughout China talking about research on deficit irrigation as a way to improve agricultural efficiency. He had always intended to bring his skills home. "It is very, very unusual to find someone who is so unselfish," says Davies of his former protégé.

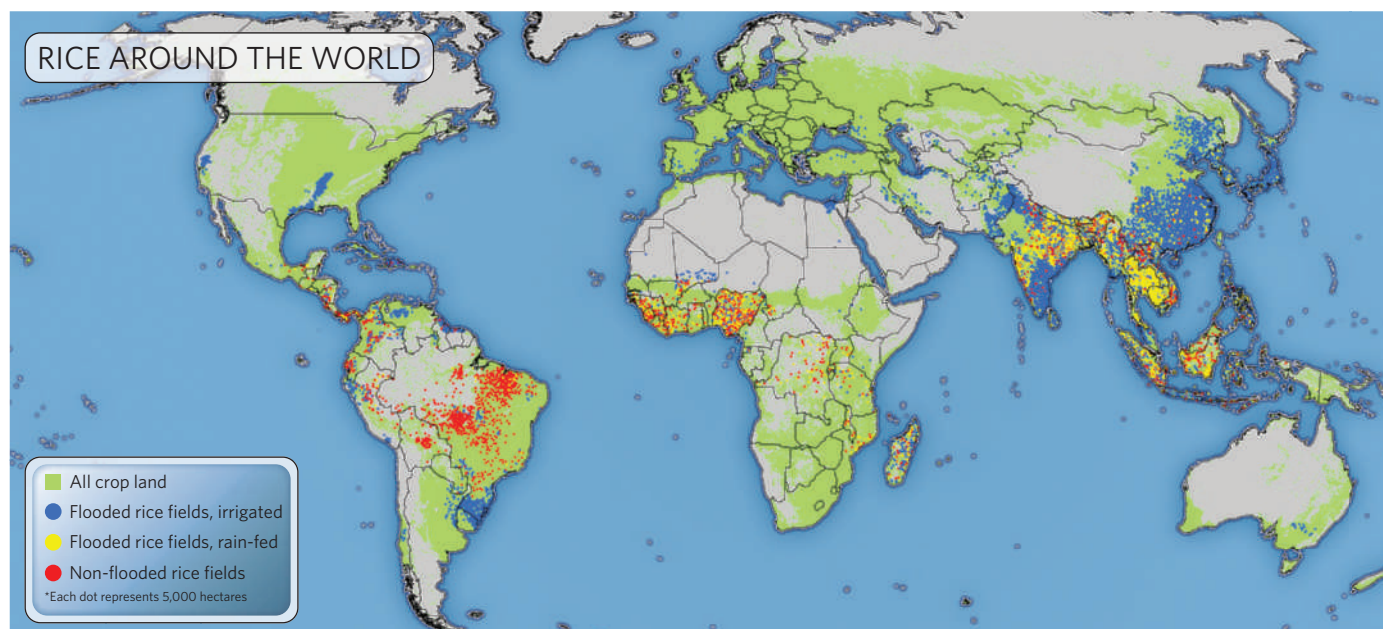
Thanks to Zhang and others, farmers in northern China have learned to use less water. In northwest China, for example, the amount of water used for irrigation has almost halved from what it was a decade ago, according to Zhang. "This is a huge and significant achievement and means that we are using less underground water, which has been depleted rapidly in recent years,"

he says. "In north China where most of the country's wheat is produced, irrigation times have been reduced from traditionally 4–5 times per crop to today's 1–2 times per crop."

"Life for peasants in the countryside is so hard," Zhang says. "We suffered a lot." He and his parents are happier now, but many of his childhood friends are still in poverty. "If you ask how much has changed in the farmers' lives, it is very little," he says. "My dream is to see the farmers have a better life and to have equal rights as city people. I always want to do something for the village people, for the peasants. I am one of them."



R. HUMANS/IRRI



The biotech humanitarian Richard Sayre

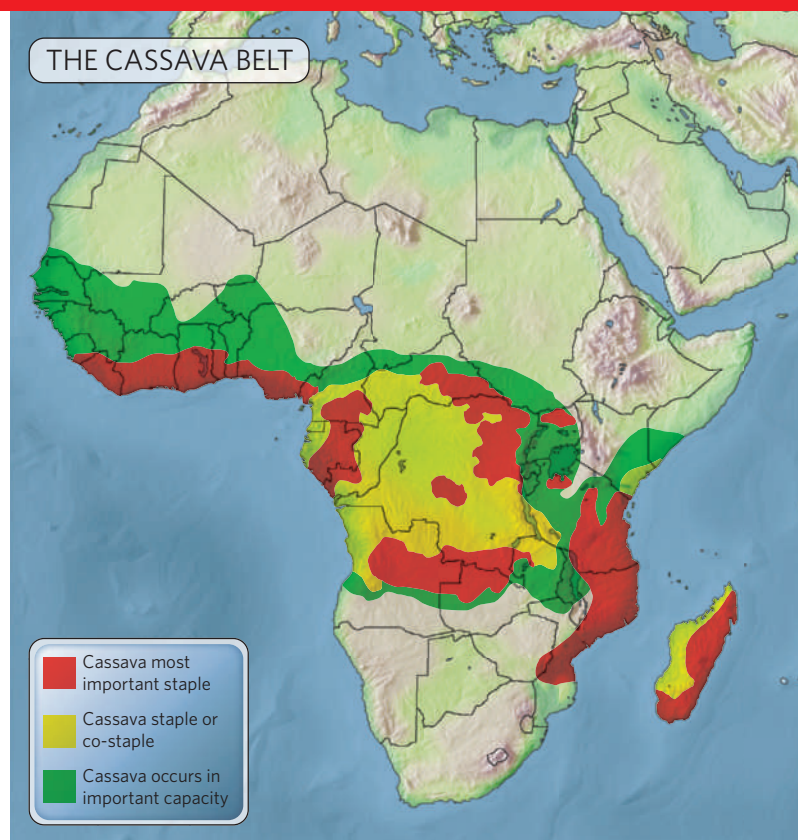
Director of the Enterprise Rent-A-Car Institute for Renewable Fuels at the Donald Danforth Plant Science Center in St Louis, Missouri
Timescale for change: 6 years

"I like being challenged," says Richard Sayre. "I couldn't do one project; I would get bored."

Sayre recently left Ohio State University in Columbus with a caravan of lab members and their families to direct the Enterprise Rent-A-Car Institute for Renewable Fuels at the Donald Danforth Plant Science Center in Missouri. In that position he will be working on turning algae into biofuel. He has also got a hand in a start-up firm called Phycal, based in Cleveland, Ohio, that works on renewable energy. And he is the head of the BioCassava Plus collaboration. Funded by the Bill & Melinda Gates Foundation, BioCassava Plus is a US\$12-million, five-year effort to turn cassava, a South American plant widely cultivated in Africa (see map, right), into a super-food that provides extra protein and a wide range of micronutrients. In his cassava role, Sayre coordinates 19 investigators from five continents and works on getting the intellectual rights to use the relevant genes as well as the tools needed to get the modified root into the public domain. He's still active in the lab, too, mostly as a troubleshooter. "Instruments break down and protocols go wrong and I'll go in and fix them," he says. "I'll go under the centrifuge and fix the brushes."

Cassava is a staple food eaten by some 250 million sub-Saharan Africans. It grows well in very poor soils and requires very little labour, so even people sick with AIDS can grow it. It also tolerates drought better than maize (corn) and other crops do. Unfortunately, it is not a very rich source of nutrients, and many varieties are tainted with bitter compounds that turn to toxic cyanide during digestion. These varieties must be soaked for a long time or ground into flour and then cooked to remove the toxins. Although this makes the food much more labour intensive, Sayre says that many farmers prefer the more-toxic varieties because thieves will not bother with roots that need such laborious preparation.

BioCassava Plus aims to improve Cassava both as a crop and as a foodstuff; so far it has managed to hit almost all the ambitious targets the collaboration promised the Gates Foundation when the programme was set up in 2005. The consortium has developed a number of genetic modifications aimed at getting the



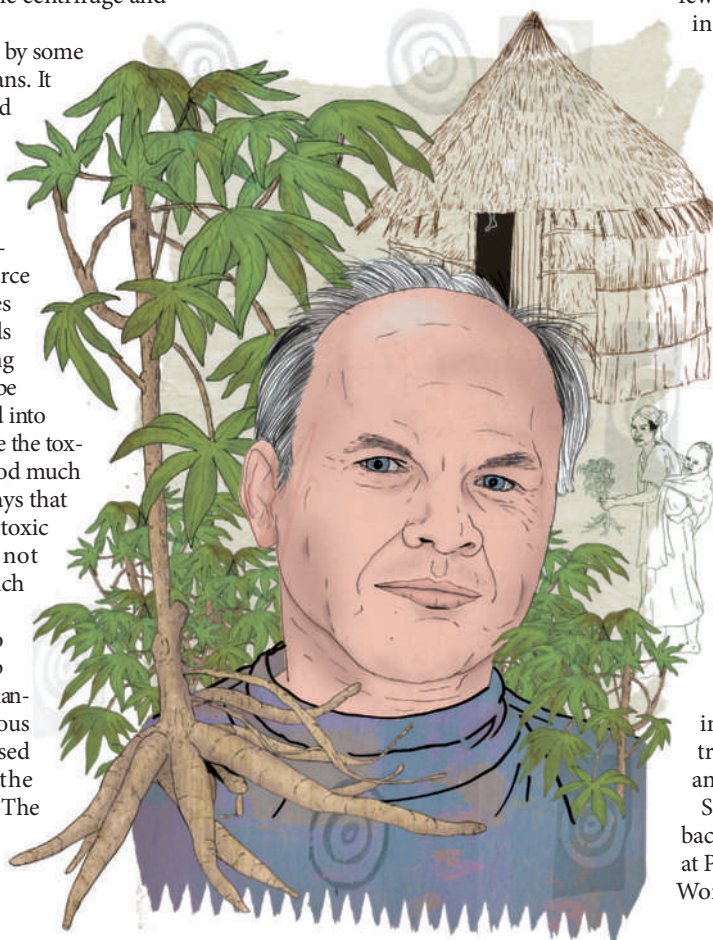
SOURCE: OKIGBO, 1980

levels of specific nutrients high enough that 500 grams of cassava would contain the minimum daily allowance. These target nutrients include protein, vitamin A, vitamin E, bioavailable zinc and iron. Another engineered strain now resists some viral infections, and yet another has many

fewer cyanogens. Many of these lines are in field trials in Puerto Rico. Next, the group will try to squeeze several or all of these traits into one, farmer-preferred variety of cassava. Sayre says they can physically deliver four genes at a time to the cassava genome, but he is not yet sure if they will all be properly expressed. Eventually, they may have to smuggle 15 genes into one line, a complex undertaking.

"All the collaborators in the cassava project are committed to changing the world, and we think we can," says Sayre. He insists that genetic manipulation is essential, despite the resistance of many African countries. "We cannot get to target levels of iron or zinc without transgenics," he says. Sayre is working to manage the resistance in his target markets by building local involvement into the scheme. "Our strategy is to have gene introgression happen in Africa by Africans," he says. "We think it is a critical element in adoption of the programme." Field trials are scheduled for Nigeria for 2009 and are being developed for Kenya.

Sayre's interest in cassava can be traced back to Offiong Mkpang, who now teaches at Palm Beach Community College in Lake Worth, Florida. In the late 1980s, Sayre



heard about a Nigerian student who needed a job. "I hired this guy as a dishwasher," he says. Sayre had a policy at the time that he would support any student or technician who wanted to work on an independent research project. "After about three months he said he would like to remove the cyanide from cassava," remembers Sayre. "He grew up in the Niger Delta; during the Biafran civil war, cassava kept his family alive." Mkpog and Sayre worked together on the project, starting with enzyme kinetics. When Mkpog moved on, Sayre kept plugging away, eventually engineering strains³ of the plant that produce 99% less cyanogen than typical cultivars.

With so much on his plate and an avowed love of stress, Sayre sounds like he just might be the uptight type. Not at all, according to Washington State University's John Fellman, one of the investigators at BioCassava Plus. "That's the paradox," says Fellman. "Most of these people who are such hard-chargers don't have time for you, but he always does." Fellman has a story that illustrates Sayre's generosity and what Fellman calls his "California laid-back" personality.

"When we were coming back from a meeting in Kampala, Uganda, we had to do some time in Amsterdam waiting for flights," says Fellman. "One of the marketing guys for the Danforth Plant Center had a world club membership for KLM. So we were sitting around drinking Heinekens at 7:30 in the morning. I was complaining I didn't have a lecture ready for my class. He said, 'here, just take this summary lecture I gave at the meeting, and this could be your lecture for Monday'. He gave me his whole presentation — I put it on my thumb drive — and every one of the slides had a specific credit on it. He wasn't saying this is my work. He was saying this is our work." ■

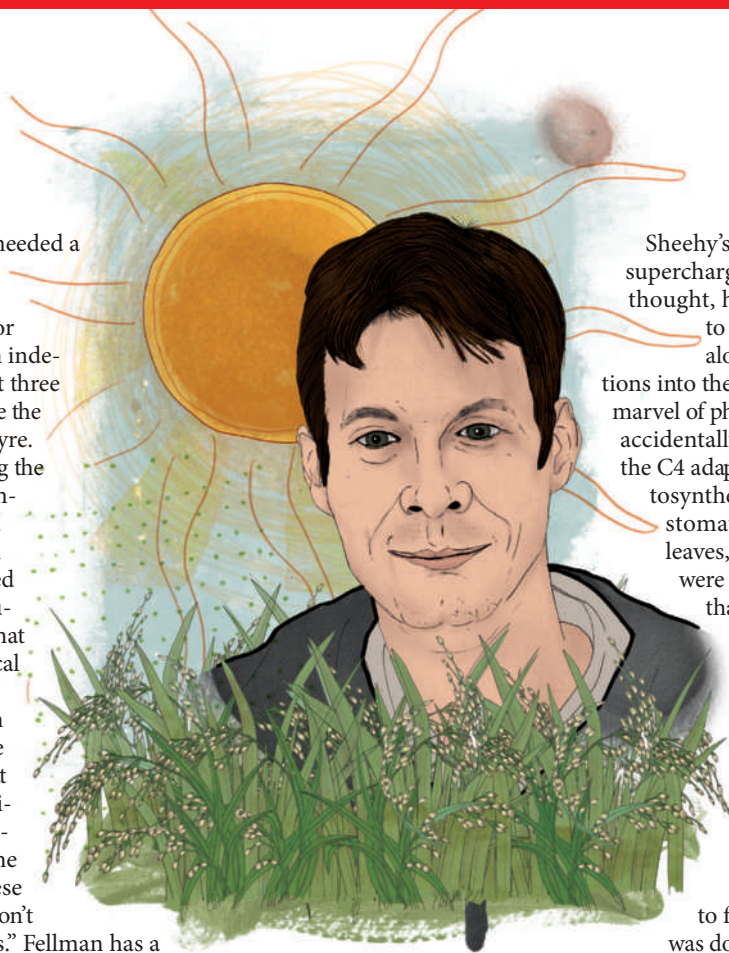
The rice transformer Julian Hibberd

Molecular biologist at the University of Cambridge, UK
Timescale for change: 15–20 years

If there was a moment when Julian Hibberd's pure science got a mission, it was halfway through a lap of the guest-house pool of the International Rice Research Institute (IRRI) in Laguna in the Philippines.

It was 2006. Hibberd was attending a meeting of a dozen scientists who study photosynthesis in rice. They had been gathered there by IRRI scientist John Sheehy, who had a very ambitious plan and wanted them on board.

Some plants, especially various grasses growing in hot climates, have evolved a way to supercharge photosynthesis. They do this by fixing CO₂ into a four-carbon sugar before feeding it into the cells doing the photosynthesizing, thus increasing the concentration of CO₂ and the efficiency of the photosynthesis. The process is known as C4 photosynthesis because of the four carbons in that sugar.



Sheehy's idea was to engineer the C4 supercharger into rice. In one stroke, he thought, he could increase yield by up to 50%. Hibberd had been asked along because his basic investigations into the still-imperfectly understood marvel of photosynthesis had more or less accidentally taken him near the heart of the C4 adaptation. While looking at photosynthesis in cells far away from the stomata that let CO₂ into a plant's leaves, Hibberd had found that they were using a number of the proteins that C4 plants use.

And so he found himself, before dinner one night, in the pool with "all the big names of C4". He remembers "swimming up and down thinking about C4 rice", with the vast experimental rice fields of the institute in the distance. "I think the most exciting thing was just to feel that the pure science that I was doing might have a route out into agriculture and make a difference to people."

Hibberd returned to Cambridge with a mission. After a lean couple of years, the Bill & Melinda Gates Foundation has recently given Sheehy's project \$11 million over three years, according to the IRRI. Meanwhile, Hibberd has been moving forward, examining which genes change as C4 metabolism evolves — as it has done on many independent occasions. "C4 leaves have modifications in biochemistry, anatomy and organelle structure," says Hibberd. "That is why it is really complicated." On just the gross anatomy level, many more veins must be made to grow in the leaves.

But there is hope. The fact that the C4 process evolved independently dozens of times means that "there is a biological precedent to say there is some relatively easy or tractable route to get these changes", says Hibberd. In particular, Hibberd's lab is looking at integrating genes from maize — which benefits naturally from the C4 pathway — into rice. Luckily, they often seem to express themselves in the right cells. "We almost certainly need to use genes from other plants," says Hibberd. "To our knowledge, there just isn't enough variation in rice to get there through conventional breeding."

Hibberd, whom fellow Cambridge plant-man David Baulcombe calls "a very modest guy", seems undaunted by the scale of the project. "When people say to me, 'don't you think that's ridiculous, to make C4 rice?' I say that I've got 30 years before I retire. It would be defeatist of me to think I can't understand the pathway pretty well. The projected benefits of having C4 rice are huge. If C4 can have 50% more yield, it would impact billions of people." ■

Emma Marris writes for *Nature* from Columbia, Missouri.

1. Tilman, D., Hill, J. & Lehman, C. *Science* **314**, 1598–1600 (2006).
2. Zhang, J. & Davies, W. J. *Plant Cell Environ.* **12**, 73–81 (1989).
3. Siritunga, D. & Sayre, R. *Plant Mol. Biol.* **56**, 661–669 (2004).

CORRESPONDENCE

Switch to ecological engineering would aid independence

SIR — In your News Feature ‘Is China ready for GM rice?’ (*Nature* **455**, 850–852; 2008), you consider the merits of using genetically modified (GM) crops for pest control. But don’t overlook the potential of ecological engineering, which can provide an important and undervalued approach to tackling agricultural problems.

Biological control in irrigated rice is a prime example of how increasing biodiversity can offer a key service to humans. Insect pests in rice are efficiently controlled by appropriate reduction of pesticide application together with landscape management (see, for example, M. J. Way and K. L. Heong *Bull. Ent. Res.* **84**, 567–587; 1994). This strategy has led to almost complete abolition of insecticide application in places such as the research fields of the International Rice Research Institute (IRRI) in the Philippines.

It was the strikingly successful example of biodiversity and ecosystem services management by IRRI that persuaded us to recruit the institute into a large international research network (J. Settele *et al.* *Nature* **453**, 850; 2008), with a view to developing sustainable land use on a global scale and refining ecological-engineering approaches.

Employing GM rice for pest control overlooks services available from natural resources that are not so dependent on big business and so do not undermine the independence of farmers and developing nations. Switching investment of resources from GM crops and pesticides to ecological engineering could be a more efficient long-term and low-cost strategy.

In such a well-studied crop as rice, ecological engineering could be implemented immediately. Research efforts should focus on developing

sustainable management schemes for ecosystem services for other key crops, in China and worldwide.

Josef Settele UFZ, Helmholtz Centre for Environmental Research, Theodor-Lieser-Strasse 4, 06120 Halle, Germany
e-mail: Josef.Settele@ufz.de
Jacobus Biesmeijer Institute of Integrative and Comparative Biology and Earth and Biosphere Institute, University of Leeds, Leeds LS2 9JT, UK
Riccardo Bommarco Department of Ecology, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

Offering unproven genetic tests to the public is irresponsible

SIR — Your News story ‘Alzheimer’s tests under fire’ (*Nature* **455**, 1155; 2008) highlights the fierce debate about the availability of genetic tests to the general public. The overwhelming amount of data concerning the major genetic risk factors for Alzheimer’s disease have recently been summarized: they show that various markers, although promising, are far from conclusive and are much farther from mainstream use than the controversial *APOE* test (L. Bertram and R. E. Tanzi *Nature Rev. Neurosci.* **9**, 768–778; 2008), whose flaws you describe.

During our 15 years of studying genetic risk factors for Alzheimer’s disease (see J. R. M. Oliveira *et al.* *J. Mol. Neurosci.* doi:10.1007/s12031-008-9124-0; 2008), our group has often been approached by private laboratories and colleagues interested in making *APOE* genotyping — and many other markers — available to the general public.

In our experience, non-affected family members are often interested in knowing their genotype, in the hope of preventing the disorder. But laypeople are rarely well equipped to understand the often murky

relationship between their genotype and phenotype. They may reach erroneous conclusions, make irresponsible lifestyle decisions or pursue unproven therapies if their test results seem alarming.

The core of the problem is a misunderstanding of crucial terms such as risk factor, prevention, cure and treatment, along with the false notion that medication is available to prevent the manifestation of symptoms. People who become aware of their genetic risk factors through a test with limited predictive power will be forced to grapple with a complex set of uncertainties. That is not a responsible way to offer medical information.

J. R. M. Oliveira Neuropsychiatry Department and Keizo Asami Laboratory, Federal University of Pernambuco, Recife-PE 50670-901, Brazil

Speaking up for economic-sciences modelling

SIR — In his Essay ‘Economics needs a scientific revolution’ (*Nature* **455**, 1181; 2008), Jean-Philippe Bouchaud cites the Moon landings as a remarkable achievement of physics and asks, rhetorically, what economics can offer as its flagship achievement. I submit that a society prosperous enough to pay for Moon landings is one such achievement.

In both cases, this was applied science (engineering and economic policy, rather than physics or economics *per se*), but previous research was crucial for making those applications possible. More recent NASA disasters and financial crises are signs that the existing science is not always applied — not that the existing science is useless.

Just like physicists (Bouchaud’s comparison group), economists are happy to use simple models when these will serve. But those

simple models are not all that there is to either subject.

This year’s Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel was awarded to Paul Krugman for his research on how imperfect competition affects trade and the location of economic activity. Other recent laureates have done research on behavioural economics (which Bouchaud calls a “fringe endeavour ... not taken seriously by mainstream economics”), on policy design when markets are inefficient, and on the behaviour of imperfectly informed economic agents. Contrary to Bouchaud’s implication, the models of the 1950s do not mark today’s research frontier in economics, any more than they do in physics.

The current economic crisis is a result of reckless behaviour by financial corporations and of the poor financial regulation that could not contain it. Let us remember, however, that the worst-case outcome is that average incomes in some countries might drop to levels last seen in the 1990s — when those same incomes were the highest ever experienced — and that unemployment might temporarily rise to 10% or 12%, which our grandparents would have considered unremarkable. The fact that we now see such a potential outcome as disastrous (as indeed we should) is a sign of the success of economics, not of its failure.

Jesper Stage Department of Economics, University of Gothenburg PO Box 640, 405 30 Gothenburg, Sweden
e-mail: Jesper.Stage@economics.gu.se

Readers are welcome to comment at <http://tinyurl.com/6y5ny3> or at <http://tinyurl.com/6gjjg86>.

Contributions to this page may be submitted to correspondence @nature.com. Please see the Guide to Authors at <http://tinyurl.com/373jsv>. Published contributions are edited.

COMMENTARY

Make secondary education universal

The time is right to push global learning beyond primary-school level, says **Joel E. Cohen**. The benefits could include a dramatically smaller increase in world population by 2050.

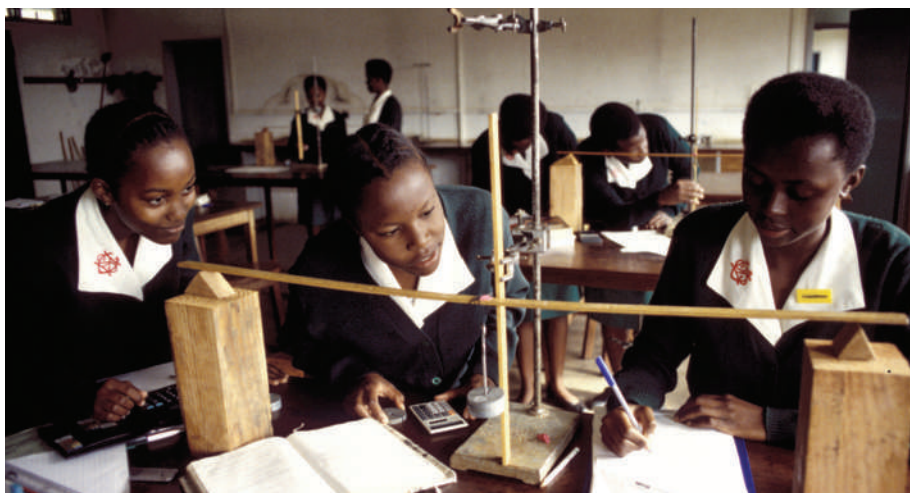
Next week, the representatives of the countries that fund the World Bank's programmes to promote universal education in developing regions will hold a planning meeting in Oslo. These leaders, and the developing countries they support, should make global secondary education a major objective.

Social, political and economic arguments for the importance of education to societies and individuals are well known¹. For example, in developing countries, every additional year of schooling increases an individual's income, on average, by 10% or more. Secondary education is a minimum prerequisite for creating the new technologies that developing countries need. These are good reasons to provide access to high-quality education for at least 10–12 years. But other demographic and environmental arguments deserve far more public attention than they currently receive.

Secondary education increases people's capacity and motivation to reduce their own fertility, improve the survival of their children, and care for their own and their families' health. Education promotes a shift from the quantity of children in favour of the quality of children. This transition reduces the future number of people using environmental resources and enhances the capacity of individuals and societies to cope with environmental change.

Earth's population is growing. The United Nations Population Division projects, as a medium scenario, that there will be 9.2 billion people by 2050 if fertility continues to follow its current declining trend (Fig. 1). If women have, on average, half a child more, or half a child less, per lifetime than assumed in that projection, world population in 2050 could be as high as 10.8 billion or as low as 7.8 billion. Thus, a difference in fertility of a single child per woman between now and 2050 alters the 2050 estimate by 3 billion, a difference equal to the entire world population in 1960.

Secondary education has the potential to influence that outcome dramatically. Although there are other factors at work, in many developing countries, women who complete secondary school average at least one child fewer per lifetime than women who complete primary school only² (Fig. 2). In Niger in 1998, for example, women who completed secondary education had 31% fewer children (on average, 4.6



R. GILING/STILL PICTURES

per lifetime) than those who completed only primary education (6.7). In Yemen in 1997, women who completed primary school had 4.6 children on average whereas women who completed secondary school had 3.1 children on average. In some sub-Saharan African societies, lifetime fertility is reduced only among girls who have had 10 or more years of schooling³. But different societies have different thresholds for reaping the demographic benefits of reduced fertility and lower population growth.

Promisingly, a woman's fertility decisions are influenced not just by her own education level, but also by that of the reproductive-age women around her. In one study, for clusters of 25 women of reproductive age in 22 countries of sub-Saharan Africa, higher average levels

of education for the women around a given individual reduced her birth rate, independent of urbanization and her individual education level⁴. Ignoring this effect, if one were to increase education from the region average level of 3.8 years of schooling to that of, say, Kenya (6.9 years) it would lower the average total fertility rate for these countries by 0.52 children. Taking into account the effect provided by educated clusters of women lowers the predicted fertility by nearly twice that. Again, education of the surrounding community reduces the total fertility rate only when the average length of education reaches four or more years, so minimum thresholds of education are required to reap the benefits of an educated social milieu.

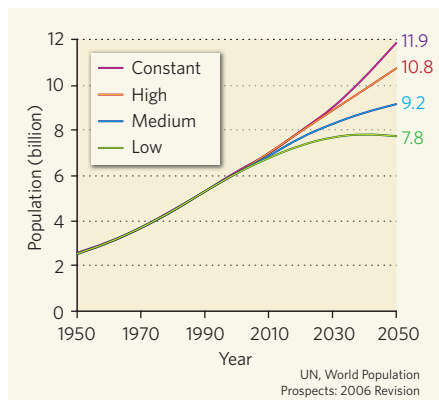


Figure 1 | Population estimates. A difference of one child per woman from now until 2050 could mean a difference of 3 billion to the global population.

Focus on the children

One of the mechanisms by which education leads to lower fertility is likely to be reductions in child mortality. Parents who are more confident of the survival of their children have reduced incentives to have many children. And the community's level of education contributes to improved survival in addition to influencing fertility⁵. Moreover, because better education in youth is associated with better health in old age, an investment in educating young people will yield reduced disability as populations age⁶.

Achieving universal secondary education will be difficult. Worldwide, among young people aged 15–24 years in 2000, more than one-half completed primary school but fewer than one-quarter completed secondary school¹.

Fewer girls than boys get secondary schooling. In 2000–04, the percentages of girls and boys in the appropriate age groups who were enrolled in secondary school were, in Benin, 17% and 38%; in Ethiopia, 16% and 28%; in Yemen, 29% and 65%; in India, 47% and 58%, respectively⁷. The deficits in schooling are concentrated in south Asia and sub-Saharan Africa. By contrast, in the more developed countries, primary completion rate exceeded 80% and secondary completion rate was close to half.

The challenges are set to increase. The number of children aged 5–14 years, all of whom will need secondary education, is predicted to rise from 1.2 billion in 2000 to a peak of less than 1.3 billion around 2025 and then decline gradually⁸ to 1.2 billion by 2050. This relatively stable number conceals important differences. During this half century, the number of children aged 5–14 years will grow dramatically in the least-developed countries — with the least means to fund education — whereas the numbers of children will be steady or declining almost everywhere else.

Pinning down costs

The costs of providing secondary education are difficult to estimate for three reasons. First, the cost per child currently in school differs from that for children not currently in school, who are more likely to be in remote locations, poor, of minority cultures and disadvantaged. Second, the quality of some schools is so inadequate that, even if education is available, parents may choose not to send their children. Third, the means of education are changing, moving from conventional schools to include information centres, homes and places of work. The costs of these and other unconventional means of education are unknown.

In developing countries, the average cost per child per year of providing secondary schooling to all eligible children is estimated at \$198 (in constant 2002 US dollars). In low-income countries, the average annual cost per child is \$125, whereas it is \$912 in the middle-income countries¹.

The aggregate annual costs of universal secondary schooling depend on how quickly universal enrolment is to be achieved and on the repetition rate — the proportion of students who repeat a grade each year. The estimated cost under present repetition rates (which average about 9% but vary widely) to achieve universal

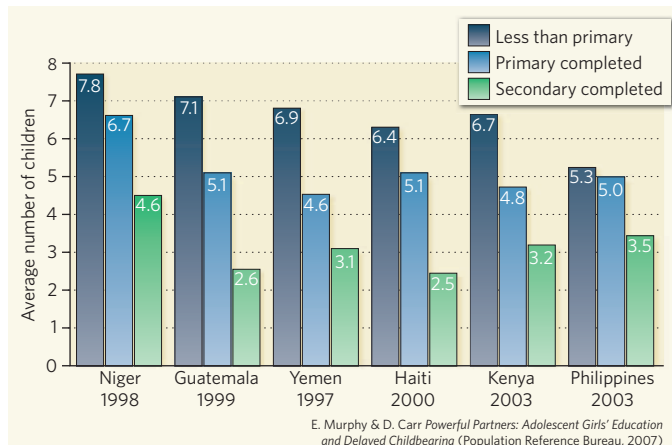


Figure 2 | Education and children. In many countries at early stages of economic development, women who complete secondary school average at least one child fewer per lifetime than women who complete primary school only. The absolute levels of fertility vary widely from country to country.

secondary education immediately would be \$62 billion annually. Although unrealistic, the figure provides a useful upper limit. If the expansion of a full secondary education to 90% of children worldwide is stretched out over 25 years and repetition is reduced to 7%, then the average annual costs for the first five years are \$5.4 billion. In this scenario, much of a generation would miss the benefits of secondary education and the demographic impacts would be smaller. Shrinking the timeline to 15 years with no change in present levels of repetition, the average annual cost for the first five years would be \$12.5 billion, and that number would go up in the remaining ten years. Whether the world can afford additional annual costs of up to \$70 billion depends on who pays. That amount would be approximately 7% of the total gross national income of all low-income countries but less than 0.3% of that of all high-income countries.

Cost is not the only obstacle⁹. Families may value more highly the time children spend working for income or doing chores. Education competes for scarce national resources with roads, medical care and defence. Returns on investment in education are difficult to measure persuasively.

Internationally comparable data on the quality of primary and secondary education are scarce or absent, so incentives to provide education of high quality are limited. The benefits of schooling accrue too slowly to boost political incumbents. Political violence disrupts schooling. Cultural discrimination inhibits schooling for girls and minorities (linguistic, religious or ethnic). The history of education in a country affects the success of externally proposed solutions to educational deficits.

To prepare to deal with these challenges, the international community needs to make a commitment now to high-quality secondary education for all. This political vision needs to be backed with more money and higher priority for education from donor and developing countries. Educators need to improve effectiveness and economic efficiency of education by using data on what and how children learn and on alternative pedagogy and technologies. Case studies of countries that perform well with the funding available to them are needed, region by region. The diversity of educational systems in different countries needs to be recognized in aid policies and educational assessments. And

the goals of primary and secondary education, whether to foster nationalism and religious fundamentalism or to embrace diversity, need to be discussed nationally and internationally¹⁰.

For the next several decades, virtually all rises in numbers of people will occur in the cities of the less-developed regions⁸. This vast shift presents an opportunity: it will be easier to reach the added children, and to attract and retain their teachers. Universal, high-quality primary and secondary education is achievable within 25 years. Educating all children well is a worthwhile, affordable and achievable strategy to develop people who can cope with problems foreseen and unforeseen.

Joel E. Cohen is the Abby Rockefeller Mauzé Professor of Populations at the Rockefeller University, 1230 York Avenue, New York 10065, USA, and professor of populations at Columbia University in New York.
e-mail: cohen@rockefeller.edu

"Universal, high-quality primary and secondary education is achievable within 25 years."

1. Cohen, J. E., Bloom, D. E. & Malin, M. B. (eds) *Educating All Children: A Global Agenda* (MIT Press, 2007).
2. Population Reference Bureau *Women's Education and Family Size in Selected Countries, 1990s* (2008). Available at www.prb.org/Educators/TeachersGuides/HumanPopulation/Women.aspx
3. Jejeebhoy, S. J. *Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries* (Clarendon Press, 1995).
4. Kravdal, Ø. *Demography* **39**, 233–250 (2002).
5. Kravdal, Ø. *Pop. Stud.* **58**, 177–192 (2004).
6. Cohen, J. E. *Sci. Am.* **293**, 48–55 (2005).
7. Ashford, L., Clifton, D. & Kaneda, T. *The World's Youth Data Sheet 2006* (Population Reference Bureau, 2006).
8. United Nations Population Division, *World Population Prospects: The 2006 Revision*. Available at <http://esa.un.org/unpp>
9. Cohen, J. E. & Bloom, D. *Finance and Development* **42**, 8–14 (International Monetary Fund, 2005).
10. Cohen, J. E. & Malin, M. B. *International Perspectives on the Goals of Universal Basic & Secondary Education* (Routledge, in the press).

To discuss this topic, visit <http://tinyurl.com/6fqgfd>.

BOOKS & ARTS

Preparing for pandemics

More deadly than the First World War, the global outbreak of influenza in 1918 terrified populations and tested governments. But would we fare any better today, asks **Michael Sargent**?

Living with Enza: The Forgotten Story of Britain and the Great Flu Pandemic of 1918

by Mark Honigsbaum

Palgrave Macmillan: 2008. 256 pp.

£15.99, \$45.00

The great influenza pandemic of 1918 was perhaps the severest test faced by public-health administrations during their first 50 years of existence. With 50 million deaths worldwide, the loss of life caused by the pandemic exceeded that in the First World War by a factor of ten. The health authorities were powerless; drugs that could halt an infection were unavailable, and preventive immunization was still a dream. Influenza's arrival was dramatic and frightening — victims collapsed suddenly, overcome with lassitude, high fevers and excruciating headaches. Those who died turned purple, drowned by fluid in their lungs generated by a fierce reaction to the virus or a secondary infection. A third of some communities succumbed to influenza, sometimes with a 5% mortality rate that fell disproportionately on young adults.

Journalist Mark Honigsbaum's new book offers a vivid account of the pandemic. It reflects on what could have been done to avert the disaster and what should be done about future pandemics. In the United Kingdom, the greatest responsibility lay with the chief medical officer, Arthur Newsholme. At a poignant emergency summit at the Royal Society of Medicine in London two days after Armistice Day, when the second and deadliest wave of influenza was at its height, he admitted that he knew of "no measure that could resist the influenza". Most doctors would have agreed. The public wanted more of their reassuring presence, but this was impossible given that medics were still needed for the war in France. Face masks and restrictions on places of public entertainment were tried as measures for containing the spread of the disease, and were adopted by some local authorities in the United States, but had little effect. Restrictions on public mobility were scarcely practical when the armed forces began to demobilize.

Important lessons came out of the emergency. Victims of the first wave were sometimes infected again in the second, suggesting that the virus could evade the immune response. On remote Pacific islands, communities with no previous exposure to influenza faced annihilation if they were not protected by quarantine.



BETTMANN/CORBIS

Children from families stricken by influenza were fed by volunteers, whose masks offered little protection.

The biomedical establishment has acquired a profound knowledge of influenza. We now know that it is caused by a single-stranded RNA virus that is extraordinarily mutable and able to alter its surface coat proteins to evade any acquired immunity. In a remarkable piece of microbial archaeology, scientists have reconstructed most of the 1918 virus from pathology specimens stored during the pandemic, showing that it resembles strains found in east Asian ducks. This was not a complete surprise: the two other twentieth-century influenza pandemics in 1957 and 1968 had similar origins, as does the H5N1 'bird flu' strain that occasionally kills humans. This leap across the species barrier seems to occur readily in the rice paddies of east Asia, where wild ducks are never far from human habitation.

Reflecting on the possible consequences of a similar pandemic today, Honigsbaum wonders if Britain's contingency plans would enable it to manage any better than it did 90 years ago, given the demand for medical services. Large-scale absenteeism from work, he believes, would cause serious economic dislocation in a world that has become much more dependent on remote providers of power, food, transport and communications. Ignoring the global perspective with his British focus, he fails to emphasize that these problems will be shared by all nations and felt most acutely in the vast cities of the developing world in which civic

organization is precarious. He suspects that people today would respond to a pandemic with less grace and stoicism.

What have we gained since 1918? We have an unparalleled knowledge of the virus and an international microbial surveillance system that acquitted itself with panache during the 2003 epidemic of severe acute respiratory syndrome (SARS). We have antibiotics that will cure secondary bacterial infections. We also have two anti-flu drugs, Tamiflu (oseltamivir) and Relenza (zanamivir), that can reduce the intensity of the disease if taken early. Although governments are stockpiling these drugs, they are unlikely to have enough available to cover entire populations; this could perhaps be rectified in a short time if a pandemic seemed imminent.

The most important preparation is continued research — for drugs to control infection and severe reactions to influenza, and for the development of mass immunization procedures that protect against all strains of influenza virus.

The book's title is a reminder of how children of the 1920s extracted a little gaiety from unpromising circumstances with a skipping rhyme: I had a little bird / Its name was Enza / I opened the window / And in-flu-enza. ■

Michael Sargent is at the National Institute for Medical Research, London NW7 1AA, UK, and the author of *Biomedicine and the Human Condition*. e-mail: msargent@nimr.mrc.ac.uk

Crimes in the name of research

Das Robert Koch-Institut im Nationalsozialismus

by Annette Hinz-Wessels

Kulturverlag Kadmos: 2008. 192 pp.
€22.50 (in German)

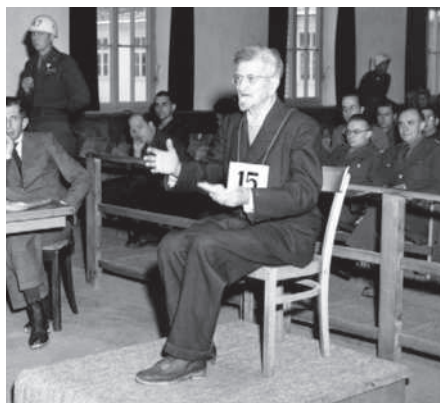
The Robert Koch Institute in Berlin was founded in 1891 and conducts research into infectious bacteria and viruses. When it celebrated its centenary, the crimes committed by members of the institute between 1933 and 1945 were apparently not of interest, and were not mentioned. Ten years later, after the Max Planck Society and the DFG, Germany's main research-funding agency, had investigated their own histories, this changed. Scientist Annette Hinz-Wessels has written the first history of the institute, concentrating on the years under National Socialism.

Das Robert Koch-Institut im Nationalsozialismus (The Robert Koch Institute Under National Socialism) begins by presenting the history of the institute from 1891 to 1933. When the Nazis took power in 1933, it had one director and six group leaders, all non-Jewish, and the reigning powers saw no need to fire any of them. But they were all of retirement age, so the intellectual power of the institute was reconstructed after 1933 according to the wishes of the Nazis.

Twelve research scientists, including postdocs and technicians, were fired because they were Jewish. Any reader would like to know what happened to these people during and after the Nazi era, yet the book says little about their fate. Was there an attempt to hire them again? One of them who survived in Berlin was invited back,

but there is no further mention of the others.

Three of the institute scientists discussed in the book — Claus Schilling, Gerhard Rose and Eugen Haagen — have been analysed previously. Schilling was one of the group leaders who retired. He had once been a state medical doctor in Togo, west Africa, where he became fascinated by malaria. After his retirement, he continued to try to construct a vaccine for the disease, and between 1942 and 1945, he used prisoners from the concentration camp at Dachau in southern Germany for his malaria experiments. Of the



Despite facing a death sentence, Schilling asked to publish the results of his unethical malaria studies.

1,200 people he infected with malaria, between 300 and 400 died. Schilling was caught by the Allies and executed in 1946.

Schilling's successor as group leader was Rose, who had worked in China and was interested in typhus. He too used prisoners from concentration camps and killed many people in his experiments. He was condemned to life imprisonment

at the Nuremberg trials after the Second World War, but was released in 1955. According to journalist Ernst Klee, who exposed many medical crimes of the Nazis in his book *Das Personenlexikon zum Dritten Reich* (Fischer; 2003), Rose became a member of the Max Planck Society in 1962.

Hinz-Wessels also discusses Haagen, who was a virologist. He was condemned to a jail sentence by a French court for his murderous experiments in a concentration camp in Alsace, in what is now eastern France. Released after a short time, he joined a virus laboratory in Tübingen, Germany, and received grant support from the DFG in 1957. Hinz-Wessels quotes from a letter Haagen wrote in August 1943 to Rose, which reads: "I contacted the central office of the SS [the Nazi protective squadron] to receive sufficient human material from worthless lives for our purpose". Schilling, Rose and Haagen are the worst scientists described and, although their experiments were already broadly known, this book adds new detail.

The book is well written and accurate. Some material might have been better presented in tables, such as the names of people fired from the institute in 1933, the names of the group leaders and those who were murdered. Published in German, it will be of interest to those now working in the Robert Koch Institute or at similar institutions. An English translation is warranted and will attract readers.

What is the bottom line? Do not try to write the history of bad scientists when any of them are still alive. Like Hinz-Wessels, do it only when all of those who were involved are dead. ■

Benno Müller-Hill is emeritus professor of genetics at the University of Cologne, 50674 Cologne, Germany. He is the author of *Murderous Science*. e-mail: muellerhill@uni-koeln.de

Conspiracy at the bench

Experimental Heart: A Novel

by Jennifer L. Rohn

Cold Spring Harbor Laboratory Press: 2008.
364 pp. \$13.99, £7.99

The scene in *The Day of the Jackal*, in which the lone assassin carefully assembles his rifle, each step and component carefully described, is one of the most memorable in literature. Such detail tells us that the author, former serviceman and war correspondent Frederick Forsyth, really knew his stuff. More importantly, it increases the tension to an almost unbearable pitch, just before the climax when

the assassin will take a shot at his target. Here is another climax:

"The answer clicked into place with an almost physical impact. Gina had inadvertently inserted the wrong gene into her herpes virus vectors, having been led to believe that it was in fact FRIP. Clearly, Rouyle had never had any intention of using FRIP in combination with her gene therapy vectors at all.

I slammed my fist onto the opened notebook. The conclusion was inescapable. And there were now two key questions remaining. One, what was the mystery gene that had been swapped in FRIP's place? And, two, was this substitution a secret only of Rouyle's, or was Pfeiffer-DeVries

in on it as well? Or even Boyd? My mind reeled at the possible layers of conspiracy."

This scene, told by a lone researcher in his laboratory, late at night, is from cell biologist Jennifer Rohn's first novel, *Experimental Heart*. It takes place in a cancer-research centre in London, in which a biotech company called Genaxis has inveigled itself like a tumour. Cell-cyclist Andy O'Hara, a postdoc who has hitherto sublimated all romantic instincts into work, sees Genaxis virologist Gina Keyser through her lit lab window late one night, and is smitten. O'Hara, ensnared, becomes involved in Keyser's work and sinks into a quagmire of intrigue, only to be resolved when he is forced to carry out some risky detective work.

At the heart of the story, which Rohn tells well for the most part, is a mystery that turns

on a question of science, based on a fictional, if realistic, cell-signalling pathway. Along the way we witness the shifting, nomadic international fellowship of scientists, and especially of postdocs. If I have one criticism, it is that the character of O'Hara does not ring true. Modest and self-effacing, he is surrounded by beautiful, talented, wise and tolerant women who care for him even after he repeatedly spurns their advances. Like the butler Stevens in Kazuo Ishiguro's *The Remains of the Day*, he remains singularly dedicated to duty.

Rohn aims to change the way in which the reading public thinks about scientists. Novels that portray scientists as protagonists, doing science in a realistic way, are scarce. This is surprising, given that there are enough detective novels to derail the Orient Express, and airport bookstalls groan with fiction featuring

the detailed work of forensic scientists, doctors, lawyers, soldiers, sailors and, in all likelihood, candlestick makers. Why?

Two forces conspire to shut science out in the cold. The first is snobbery, in which the literary establishment has never shaken off the idea that scientists are unfit to join the high table of culture. The second is that scientists are viewed as modern-day wizards whose occult works are unintelligible and probably dangerous. People cannot make up their minds whether scientists are sources of unbelievable good ('miracle cures') or chthonic evil ('frankfoods'); so the breed is best left alone. Were scientists to show signs of literacy, intelligibility or humanity, both constituencies would ignore them, the first by braying more loudly, the second for fear of puncturing cherished prejudices. Occasionally, both forces come together — for example, in the

oft-expressed confessions of media types that they are very poor at maths or science, as if this admission were a badge of honour rather than of shame. In most circles, to confess any mathematical or scientific ability risks ostracism.

Because of these prejudices, prospective readers of *Experimental Heart* have to muster the activation energy required to surmount the barriers that society has put in the way of understanding scientists as people. The worry is that when they hear the dreaded 's' word, many won't even bother. In the same way that female scientists notoriously have to accumulate a far more stellar publication record than their male peers to achieve tenure, novels about scientists have to work twice as hard as detective stories to reach the same place in the best-seller lists. ■

Henry Gee is a senior editor at *Nature*. His latest novel is *By The Sea*.

Q&A: Electronic music comes of age

The Columbia-Princeton Electronic Music Center in New York was the first institution of its kind in the United States for experimenters seeking new technology-based sounds. Fifty years after its founding, director of research **Doug Repetto** explains how electronic music has evolved and how the role of academic music centres is changing.

Why was such a centre needed?

One of the primary reasons was equipment: when the centre was established in 1958, we had a room-sized synthesizer and giant tape decks — things that individuals couldn't have. It was a physical resource: everyone had to come here to make this music. It had influential students, including Bob Moog, who went on to make the Moog synthesizers that revolutionized the way technology is used in music around the world.

In 1996 it was renamed Columbia University's Computer Music Center. Why?

The change acknowledged that hardware was no longer the focus. Today every student has a laptop, and centres around the world have had to struggle with the question of what to do now they're not a physical resource any more. In the past decade, the centre has become less about academic computer music and more about multimedia and inter-media work.

How did you start making electronic music?

I've always been interested in music. In the beginning it was through being in rock bands. Then I went to college and discovered experimental contemporary music. But I was interested in systems and processes, almost doing experiments in music. It was difficult



Doug Repetto experiments with electronic sounds.

to convince the other college kids that they should be partaking in these experiments. They weren't that interested in playing one note for an hour or playing things very quickly or very slowly. That got frustrating until I discovered computers. Computers would do whatever you wanted.

Your own work fuses music, electronics and sculpture. What inspires you?

My work has moved away from explicitly being music. I build environments where things happen regardless of what you do. You can perturb the system and inject some

energy into it, but it still does its own thing. The inspiration comes from the idea that the physical world is out there and there are ways to understand it.

How do you think perceptions of electronic music have changed?

It's hard to find someone now who thinks electronic music is not music. For a long time, there was a requirement at Columbia University that people doing doctorates in composition had to produce a music score. But then people started doing dissertations that used computers. They would print out the source code from their software and say 'that's the score'. The university had to expand what it counts as music.

What will be happening in 50 years' time?

In a curious inversion, the music world has gone from entirely personal technology to almost entirely large-scale industrial technology, and now it's coming back to being personal again. The next 50 years will be mostly more of the same: artists pushing the boundaries, misusing technologies and asking difficult 'how' and 'why' questions about new technologies. Hopefully, we will end up thinking about things in ways we wouldn't have if artists hadn't asked those questions. ■

Interview by **Daniel Cressey**, a reporter for *Nature*.

S. FRIEDLANDER

Hidden treasures: The Jagiellonian Museum, Kraków

Poland's oldest university museum celebrates the Sun-centred ideas of Copernicus and the history of the nation itself, explains **Alison Abbott**.

Nicolaus Copernicus was, they say in Kraków, one of the city's first tourists. He 'almost certainly' visited its famous salt mines, they insist. And indeed, some 150 metres underground, among the statues of saints and heroes in the labyrinth of chapels and chambers carved out of the salt rock by miners over the centuries, stands an imposing likeness of Poland's most famous scientist — the man who worked out that Earth orbits the Sun.

Copernicus was a student in Kraków from 1491 until 1495, during one of Poland's intermittent spells of greatness in between the long periods of oppression by its neighbours Russia, Prussia, Sweden and Austria. More than a century earlier, Wladyslaw I the Elbow-high, a fine warrior who was crowned king in 1320, brought the necessary order, unity and wealth to the fragmented chaos of the Polish territories to give his son and (no doubt taller) successor, Casimir III the Great, the ability to create institutions worthy of an aspiring European power. The Kraków Academy was founded by papal decree in 1364, making it one of the oldest seats of learning in Europe.

But the academy — now known as the Jagiellonian University — only really took off in 1400, thanks to the vision of Queen Jadwiga, one of the first regents to be democratically elected from among the increasingly powerful Polish nobility. She bequeathed her wealth to the university, leaving her illiterate widowed husband to realize her plans. Uniquely for its time, its professors did not have to depend on student fees for their income. Instead, they were paid from the proceeds of the salt mines, which stopped full operation only in 1996.

Throughout the fifteenth century, the university's reputation soared. It attracted students from all over Europe, thanks in no small part to its strength in mathematics and astronomy. Copernicus must have benefited from its special intellectual atmosphere, even though a fire swept through the campus the year after he arrived, razing it to the ground. The lovely Collegium Maius — then, as now, the main seat of the university — was built on the ashes, a beautiful square of gothic buildings around a courtyard that was given a neo-Baroque makeover in the mid-nineteenth century. The college attracts

as many tourists today as the popular salt mines. It also houses the Jagiellonian Museum, which brings together scientific instruments from the university's glory days, collections put together by professors over the centuries and bequests from grateful students that continue today.

But by the mid-sixteenth century, the university's glory was waning. Polish nobles lost their noble vision

reaching its peak in the 1770s, and was brief. Students were sent abroad to observe research and bring that scientific culture back to their alma mater. But many never had a chance to return. By 1795, a series of three partitions — Russia annexing from the east, Prussia from the west and Austria from the south — had wiped Poland from the map. The country did not exist again until after the First World War. Kraków was swallowed up by Austria, and returning scholars found themselves ineligible to work at the exclusively German-speaking university.

Eighty years passed and the university began to flourish once more, at least until the First World War. One of its triumphs during that time was the first liquefaction of the components of air in a stable form, by chemist and instrument-maker Karol Olszewski and his colleague Zygmunt Wróblewski.

The disaster that was wrought by the two twentieth-century world wars is well known. But the university never closed down, and the Collegium Maius has been spruced up since the fall of the Iron Curtain. Its extensive museum includes among its countless treasures the famous Jagiellonian Globe, which shows the American continents for the first time. Made around 1515 and acquired by Brożek, the 7-centimetre-diameter globe is surrounded by an armillary sphere and includes a clock mechanism. The museum also houses a vast, valuable collection of optical instruments, including many telescopes, and Olszewski's beautifully crafted cryogenic equipment.

Copernicus never returned to Kraków after his student days. But in the 1950s, the museum commissioned the construction of three massive timber instruments of the type that Copernicus probably used — a quadrant, a triquetrum and a skeletal armillary sphere, all of which measured the position of celestial objects in the pre-telescope era. The majestic trio, which stand between 1.5 and 2.5 metres high, are as much a homage to science as the Copernicus salt statue deep inside the Earth. ■

Alison Abbott is *Nature's* senior European correspondent.

See <http://tinyurl.com/5ash6p>. For more hidden treasures, see www.nature.com/nature/focus/hiddentreasures.



The gilded Jagiellonian Globe within its armillary sphere is the first to show the Americas.

and began to use their power to promote their own self-interests. They decided that offices of the state and church should be allocated exclusively from among their number, making higher education less valuable. The university faculties became embroiled in a decades-long struggle with the Jesuits, who wanted to establish a rival academy. And at the same time, wars, epidemics and other disasters were taking their toll on the weakening Polish state. The university stagnated, although the renowned astronomer and mathematician Jan Brożek (1585–1652), worked there during this time. A committed propagator of the Copernican heliocentric theory, Brożek also collected many of the important astronomical artefacts now on view in the museum.

The Enlightenment arrived late in Poland,

ESSAY

Generosity: A winner's advice

Mathematical models can reveal how prosocial human behaviour — and even social intelligence and language — have evolved, argues **Martin A. Nowak**.

One day while I was still at Oxford, Bob May gave me some advice: “You never lose for being too generous”. I was impressed because Bob is a winner. To him winning a game is everything. He has thought more deeply about winning and losing than anyone else I know. As his wife once said, “When he plays with the dog, he plays to win.” At the time, Bob was not only my adviser but also one to the British government. A few years later he would become president of the Royal Society, Lord May of Oxford and the recipient of many prestigious awards.

A mathematical analysis of human behaviour suggests that Bob was right. Generosity is an essential feature of winning strategies in games that explore human interactions.

These strategies underpin many of the choices people make in everyday life, and shed light on how our unusually cooperative ways have evolved.

Biologists recognize two fundamental forces of evolution: mutation and selection. I want to add a third: cooperation. Cooperation occurs when one individual pays a cost so that another receives a benefit. Here, cost and benefit are measured in terms of reproductive success. Reproduction can be genetic or cultural, the latter involving the spread of knowledge and ideas.

Only if certain mechanisms are involved can natural selection favour individuals who reduce their own fitness to increase that of a competitor. One such mechanism is direct reciprocity: my strategy depends on what you have done to me. Another is indirect reciprocity: my strategy depends on what you have done to me and on what you have done to others.

In both, mathematical analysis shows that winning strategies tend to be generous, hopeful and forgiving. Generous here means not seeking to get more than one's opponent; hopeful means cooperating in the first move or in the absence of information; and forgiving means attempting to re-establish cooperation after an accidental defection. These three traits are related. If I am generous, it is easier for me to forgive, and also to be hopeful and take the risk of cooperating with newcomers.

In the Wimbledon championship, you must defeat your opponent to move to the next

round. But everyday life is not like a tennis tournament. Instead, most of our interactions occur in a population of players, and pay-off accumulates over encounters with many different people. Because overall success is proportional to that pay-off sum, the other person in any one encounter is more a partner than an opponent. If I am willing to let others have a slightly bigger share of the pie, then people will want to share pies with me. Generosity bakes successful deals.

Experiments have confirmed the success of generosity. A typical set-up involves students and computer screens. The computer pairs

but sometimes even when you have defected. In other words, I am forgiving. For a while, cooperation thrives. But in a generous tit-for-tat population, the emergence of unconditional cooperators eventually invites the invasion of defectors. This leads to cycles of cooperation and defection — which could account, at least in part, for the mix of cooperators and defectors that persists in human societies.

Mathematical models allow a precise investigation of fundamental aspects of human behaviour. The games described here occur in every society. Ancestral humans spent most of their time in small groups where interactions

were repeated. The same is true for most dealings in modern life: repeat encounters are always possible and reputation is typically at stake. The evolution of prosocial behaviour cannot be understood

outside the framework of direct or indirect reciprocity. Indeed, I believe that games of indirect reciprocity have provided the crucial selection pressures for social intelligence and language.

In such games, social intelligence is needed to monitor and interpret the interactions of others. We follow with great interest what our fellow creatures do to us and to others. When deciding how to act, we take into account — often subconsciously — the possible consequences for our own reputation. Moreover, our own observations are often not enough; we want to learn from the experiences of others. Spreading the rumours of indirect reciprocity requires language. As my colleague David Haig once remarked “for direct reciprocity you need a face, for indirect reciprocity you need a name”.

Martin A. Nowak is professor of mathematics and biology at Harvard University and director of the Program for Evolutionary Dynamics, 1 Brattle Square, Cambridge, Massachusetts 02138, USA. He is the author of *Evolutionary Dynamics*. e-mail: nowak@fas.harvard.edu

See <http://tinyurl.com/6pc5hq> for further reading. For more on Being Human, see www.nature.com/nature/focus/beinghuman.



G. BECKER

NEWS & VIEWS

CANCER STEM CELLS

Here, there, everywhere?

Connie J. Eaves

Can every tumour cell propagate human cancers or is this property exclusive to an elite subset? Findings are divided. The latest set shows that — depending on circumstances — both perspectives can be correct.

A long-standing goal of both researchers and oncologists is to establish a framework for understanding how many and which tumour cells must be eliminated for treatment to be successful. One framework that has received much attention recently attempts to understand cancers as perturbed versions of the normal tissue in which they arise, with retention of many tissue-specific developmental features. The 'cancer stem cell' hypothesis is a derivative of this framework, and states that cancer cell populations have a hierarchical developmental structure in which only a fraction of the cells — the cancer stem cells — can proliferate indefinitely. If this concept is correct, it could have implications for cancer treatment. But Quintana *et al.*¹ report on page 593 of this issue that, for at least one type of human cancer, the reality may be more complex.

What is the evidence for cancer stem cells in human tumours? The creation of mice that are sufficiently immunodeficient to tolerate the growth of primary human cells in them has only recently allowed this idea to be examined experimentally, and so far only a few tumour types have been investigated^{2–7}. All of these studies suggest that only a tiny subset of cells (0.0001–0.1%) have the ability to generate a new tumour in most immunodeficient mice that were available at the time of each study. The earlier studies also found that the features of cells with tumorigenic potential differ from those of the bulk of the tumour cells but are often shared with the normal stem cells of the same tissue. Together, these findings supported the idea that human cancer cells that can produce tumours in immunodeficient mice represent a biologically distinct set with stem-cell-like properties.

Quintana *et al.*¹ transplanted single human cancer cells into highly immunocompromised mice — more so than the animals used previously — and used rigorous procedures to measure the frequency of tumorigenic cells in different strains of mice. They demonstrate that as many as one in four tumour cells derived from a type of human skin cancer called melanoma can initiate a tumour. Moreover, tumour cells capable of producing a new tumour can have many different features, most

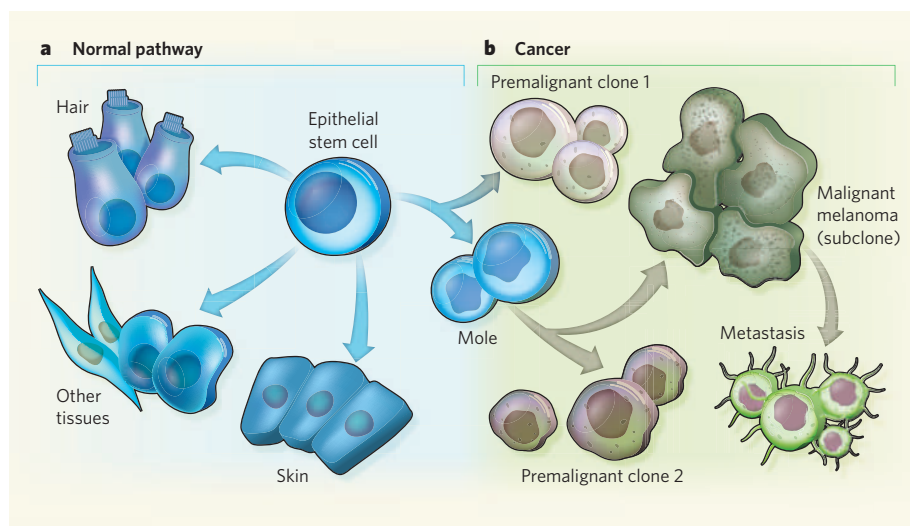


Figure 1 | Oncogenesis is a sequential process. **a**, Normally, stem cells (in the example shown, epithelial stem cells) give rise to various differentiated cells such as hair and skin cells. **b**, In cancer, however, it is thought that these stem cells must undergo a specific set of rare changes to produce a population that is frankly malignant. Such a sequential process allows many subclones to develop along the way, each characterized by more aggressive features.

of which are common to some, but not all, of the tumorigenic cells, and none of which shows a particular association with tumorigenic potential¹.

How can these surprising findings be reconciled with the earlier observations^{2–7}? One explanation could be that the process of oncogenesis typically involves an accumulation of many rare events that leads to altered control of growth and differentiation. The first of these events is likely to occur in normal stem cells because these are the only cells that undergo sufficient rounds of cell division to accumulate the particular combination of changes necessary for a malignant cell to evolve (Fig. 1). Evidence for both a clonal origin of tumours and their evolution through an accumulation of genetic mutations has been around for decades. But linking these observations to the idea that malignant cells ultimately originate from a stem-cell population is less well established.

A big impetus for the concept of cancer stem cells comes from studies of a type of blood cancer called chronic myeloid leukaemia. Here, a small abnormal chromosome called the

Philadelphia chromosome is found in several types of blood cell and in their most primitive precursors. The chronic-phase clone becomes dominant because the mechanisms controlling how many mature cells should be produced are highly deregulated, without affecting the ability of the cells to execute normal differentiation programs. If the disease is not adequately treated, more aggressive subclones, characterized by additional mutations that disrupt the differentiation process, arise, and these subclones then produce a rapidly fatal acute leukaemia. This sequence of events illustrates how premalignant stem-cell-driven cell populations may precede the appearance of a frankly malignant subclone that could contain a high proportion of cells with tumorigenic activity (Fig. 1b). The more aggressive the subclone, the faster it is likely to grow, quickly diluting the original premalignant population and making its identification difficult.

A second explanation may involve the possible influence of the environment in which the tumour cells are trying to grow. Increasing evidence suggests that non-malignant host-cell

populations have a significant role in tumour growth⁸. Quintana *et al.*¹ report that, when they used the same *in vivo* assay protocol as was used in a previous study⁷, they detected the same low (one in a million) frequency of tumorigenic human melanoma cells. Nonetheless, the authors¹ found that many more cells could form tumours when various aspects of the original protocol⁷ were altered. These included prolonging the observation period, injecting the tumour cells into an extract rich in extracellular-matrix components such as laminin to improve tumour-cell viability, and using more highly immunodeficient strains of mice as hosts.

If the growth potential of a cancer does depend on a rare subset of cancer stem cells, it seems important to know how to eradicate these particular cells. Similarly, assessing the ability of a candidate therapy to destroy these cells would seem crucial to predicting its efficacy. But even these assumptions are being challenged by some experimental findings. For example, the outcome of treating patients in the chronic phase of chronic myeloid leukaemia with the drug imatinib mesylate — before the acute phase has begun — suggests that selective and effective killing of the differentiating cells in the premalignant clone may be sufficient to prevent this cancer's progress for many years, even though the treatment has little ability to destroy the chronic-phase stem cells^{9,10}.

Research into human cancer stem cells is still gathering steam, so it is not clear how unusual or clinically relevant Quintana and colleagues' observations¹ are. It is possible that their findings are unique to a subset of tumours, to specific types of mutation, to certain states of cancer progression, to distinct factors within the tumour environment, and/or to the states of innate and acquired immunity of the host. Equally possible is that these observations are more commonly applicable. Either way, this study provokes healthy scepticism in the absolute value of any tumour-initiating cell measurement and points to the need for careful studies designed to test new biomarkers and therapeutics. ■

Connie J. Eaves is in the Terry Fox Laboratory, British Columbia Cancer Agency, 675 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada.
e-mail: ceaves@bccrc.ca

1. Quintana, E. *et al.* *Nature* **456**, 593–598 (2008).
2. Bonnet, D. & Dick, J. E. *Nature Med.* **3**, 730–737 (1997).
3. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. *Proc. Natl Acad. Sci. USA* **100**, 3983–3988 (2003).
4. Singh, S. K. *et al.* *Nature* **432**, 396–401 (2004).
5. O'Brien, C. A., Pollett, A., Gallinger, S. & Dick, J. E. *Nature* **445**, 106–110 (2007).
6. Ricci-Vitiani, L. *et al.* *Nature* **445**, 111–115 (2007).
7. Schatton, T. *et al.* *Nature* **451**, 345–349 (2008).
8. Garcion, E., Naveilhan, P., Berger, F. & Wion, D. *Cancer Lett.* doi:10.1016/j.canlet.2008.09.006 (in the press).
9. Bhatia, R. *et al.* *Blood* **101**, 4701–4707 (2003).
10. Druker, B. J. *et al.* *N. Engl. J. Med.* **355**, 2408–2417 (2006).

See also Abstracts, page xiii, and News, page 553.

MATERIALS SCIENCE

Clear leap for superconductors

Darrell G. Schlom and Charles H. Ahn

Electric fields offer an innovative means of controlling condensed-matter systems. The approach has been applied to nanoscale oxide interfaces, for studying the physics of two-dimensional superconductors.

In recent years, the ability to engineer the interface between two oxides on the nanoscale with atomic-layer precision has led to remarkable advances in our knowledge of the electronic phenomena occurring at these interfaces. For instance, the transfer of charge at the interface between two ordinary oxide insulators, LaAlO₃ and SrTiO₃ (ref. 1), gives rise to a conducting two-dimensional gas of electrons that, at low temperatures, is a superconductor (it conducts electricity without resistance)². Now Caviglia *et al.*³ (page 624 of this issue) show that an electric field can be used to modify the electron concentration of a transparent (clear) electron gas that forms at the LaAlO₃/SrTiO₃ interface (Fig. 1), allowing the emergence of two-dimensional superconductivity to be examined in unprecedented detail.

A major challenge in the study of these interfaces is identification of the source of the charge carriers (the electrons). One possible source is a 'polar catastrophe', in which the change in net layer charge (positive or negative) between the last TiO₂ layer of the SrTiO₃ structure and the first LaO layer of the LaAlO₃ structure, drives an electronic reconfiguration that introduces charge carriers into the system¹. Alternatively, they might be produced in a more mundane fashion by means of oxygen defects or the interdiffusion of cations between the LaAlO₃ and SrTiO₃ layers⁴. Caviglia and colleagues' work does not address this issue definitively, but it is not germane to the fundamental physics that they elucidate here.

Caviglia *et al.*³ use an electric field to reversibly pump charge carriers into or out of the interface, a process called charge doping. Impressively, this allows the authors to map out a region of the phase diagram of the interface state — a plot of temperature versus charge-carrier concentration that shows the phase boundary between the insulating state and the superconducting state. Because electric-field doping introduces no chemical or structural disorder at the interface, unlike chemical doping (in which the chemical composition is altered to introduce charges into the system), the authors are able to study the charge physics of a 'clean' system, without the obscuring, undesirable effects of disorder.

Interestingly, Caviglia *et al.* find no evidence of magnetism at the LaAlO₃/SrTiO₃ interface over the entire range of carrier concentrations studied, including densities of about 10¹³ electrons per cubic centimetre, for which a magnetic state was previously identified⁵.

What they find instead is a superconducting, dome-shaped region in the phase diagram (see Fig. 3 on page 625), occurring over a range of charge-doping levels and temperatures that is consistent with the range over which doped SrTiO₃ is known to be superconducting.

Caviglia and colleagues' pivotal finding is that this clean, disorder-free doping approach allows them to identify the quantum critical point that separates the superconducting state from the insulating state as a function of charge density. A detailed analysis of this critical region suggests quantum phase fluctuations that are reminiscent of insulator–superconductor quantum phase transitions, tuned by electric fields, seen in other two-dimensional systems⁶.

But where do we go from here? One clear role for the authors' approach of using an electric field to control superconductivity will be in elucidating the nature of phase transitions in other systems, particularly in systems in which the symmetry is broken by the presence of the interface. What's more, this approach will facilitate the development of nanoscale superconducting circuits, including the creation of rewritable arrays of Josephson junctions (two superconductors separated by a thin, non-superconducting region).

The authors' method will find even more wide-ranging applications. First, the electric-field control of interface properties, combined with the nearly unlimited possible combinations of oxides displaying diverse electronic properties, will allow other physical phenomena to be probed, including magnetism, orbital ordering and charge ordering⁷. The present interest in this approach is also partly due to its potential for achieving charge-mediated magnetoelectric coupling in composite multiferroic systems. These materials allow magnetism to be controlled in the solid state using electric fields. Furthermore, in its search for an eventual replacement for the transistor, the semiconductor industry is looking at exotic phenomena in various solid-state systems, including the control and manipulation of exotic ground states in novel, artificially structured materials. Electric-field control of these ground states could form the basic logic states ('0' and '1') of future information-processing technologies.

But most notably, Caviglia and colleagues' work³ marks the dawn of a new era in the design of superconductors consisting of materials created with atomic-layer precision. In these

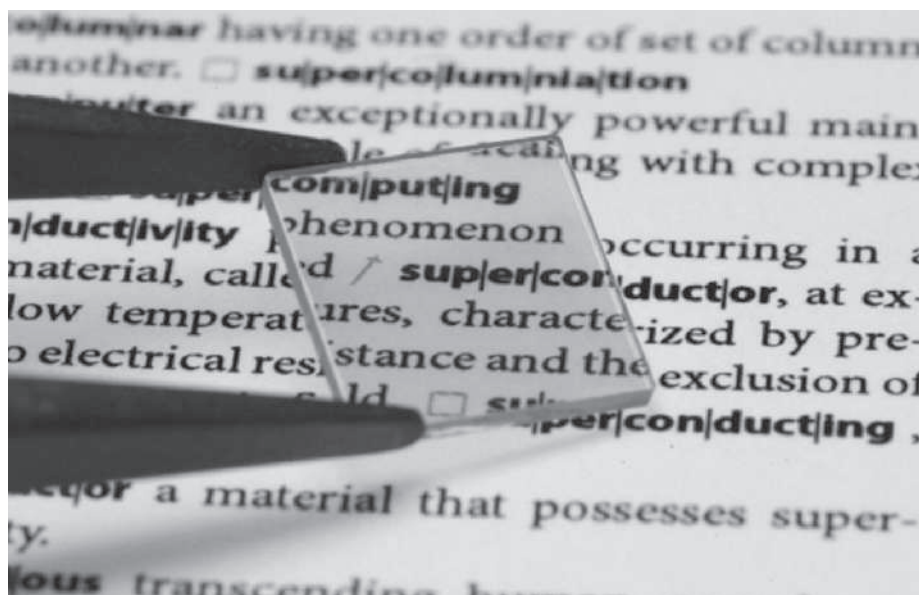


Figure 1 | Image of the $\text{LaAlO}_3/\text{SrTiO}_3$ structure. A transparent (clear) gas of electrons is formed at the atomically abrupt interface between two oxide insulators — LaAlO_3 and SrTiO_3 — laid on top of each other. Caviglia *et al.*³ use an electric field to control the electron density at this interface and thereby study the emergence of superconductivity in the electron gas. (Image courtesy J. Mannhart.)

systems, huge strains can be imparted on the constituent monolayers assembled on top of one another. This could enable electronic-reconfiguration and orbital-ordering effects to enhance the superconductivity, which could be turned on and off simply by using an electric field to modulate the charge-carrier density. Engineered oxide films, in particular, offer a new means of searching for superconductivity that traditional methods lacked. Several decades ago, it was proposed that interfaces could be used to boost the temperature at which superconductivity can exist^{8,9}, but it was also noted that this would be a challenge to achieve experimentally⁹. Studies such as that of Caviglia *et al.*, together with advances in theory¹⁰, now show that this idea is becoming a reality. ■

Darrell G. Schlom is in the Department of Materials Science and Engineering, Cornell University, Ithaca, New York 14853-1501, USA. Charles H. Ahn is in the Departments of Applied Physics and Physics, Yale University, New Haven, Connecticut 06520-8284, USA. e-mails: charles.ahn@yale.edu; schlom@cornell.edu

1. Ohtomo, A. & Hwang, H. Y. *Nature* **427**, 423–426 (2004).
2. Reyren, N. *et al.* *Science* **317**, 1196–1199 (2007).
3. Caviglia, A. D. *et al.* *Nature* **456**, 624–627 (2008).
4. Eckstein, J. N. *Nature Mater.* **6**, 473–474 (2007).
5. Brinkman, A. *et al.* *Nature Mater.* **6**, 493–496 (2007).
6. Pareno, K. A. *et al.* *Phys. Rev. Lett.* **94**, 197004 (2005).
7. Ahn, C. H., Triscone, J.-M. & Mannhart, J. *Nature* **424**, 1015–1018 (2003).
8. Ginzburg, V. L. *Phys. Lett.* **13**, 101–102 (1964).
9. Allender, D. *et al.* *Phys. Rev. B* **7**, 1020–1029 (1973).
10. Chaloupka, J. & Khaliullin, G. *Phys. Rev. Lett.* **100**, 016404 (2008).

STEM CELLS

Makeshift sperm production

Allan C. Spradling

Early middle age is a difficult time, not least for male fruitflies when sperm production falls. The unexpected reason for this decline seems to be that, as tissues age, maintaining functional stem cells becomes difficult.

Stem cells arose early in evolution and underlie the body plans of most multicellular animals. So we tend to picture stem cells and their niches — the microenvironments in which they operate — as finely tuned systems with near perfect, almost magical, capabilities. A report by Yamashita and colleagues (Cheng *et al.*, page 599 of this issue)¹ reminds us,

however, that stem cells are no exception to François Jacob's description² of evolution as "tinkering".

The subject of the Yamashita group's studies is the germline stem cells (GSCs) in the testis of adult male fruitflies (*Drosophila melanogaster*). The authors discovered that by day 20 of adulthood — early middle age for a fruitfly — a

significant fraction of GSCs had become arrested; that is, they had ceased the cell-cycle activity and asymmetric cell division that lead to each GSC producing one self-renewing cell and eventually one differentiated sperm cell. This slows sperm production without changing stem-cell number.

Normally, GSCs align one of their two centrosomes — key organelles controlling the direction of division — with the niche, which in this case is the periphery of a cluster of support cells called the hub (Fig. 1a, overleaf). This arrangement ensures that, on division, one daughter remains a stem cell while the other one differentiates. Cheng *et al.*¹ found that centrosomes in the arrested stem cells had become misaligned. Remarkably, it seems that centrosome misalignment is a by-product of the dedifferentiation of older germ cells that enter the niche to replace GSCs that have turned over; dedifferentiation randomizes centrosome orientation. These observations provide a much clearer picture of the ageing *Drosophila* testis, and hint at processes that may limit cell production in many other tissues: stem-cell homeostatic mechanisms are not perfect, and cannot maintain full activity even into middle age.

A case in point is the signature ability of stem cells to divide asymmetrically — to self-renew. Centrosome alignment seems to be a highly adaptive mechanism that acts to ensure self-renewal by limiting stem-cell loss or overproliferation³. Keeping one centrosome, the maternal centrosome⁴, permanently localized within the cell cytoplasm adjacent to the boundary between the GSCs and the niche-generating hub ensures that divisions will be oriented perpendicular to the niche (Fig. 1a). Because the niche microenvironment that maintains the stem-cell state exists only in a one-cell-wide strip around the hub, such an orientation guarantees that just one GSC and one differentiating germ cell will be produced by each division (Fig. 1b). This elegant system, or a variant of it, seems to be used in other stem cells, such as *Drosophila* neuroblasts^{5,6}. It may have evolved from a fundamental molecular asymmetry in all dividing cells — the non-equivalence of newly duplicated centrosomes — and hence may be an ancient stem-cell mechanism.

We must now ponder the fact that, despite the operation of this system, many GSCs are lost with age anyway and have to be replaced. Moreover, replacement cells seem to have difficulty realigning with the niche, and during realignment they cannot function as GSCs (Fig. 1c). However, it is still too early to blame these problems on centrosome misalignment as such. Age-related changes in niche adhesion or signal reception could underlie stem-cell loss, and dedifferentiation itself might be the slow step in replacement and simply be a prerequisite to alignment. Nonetheless, the system as a whole comes under stress by early middle age, and problems only get worse in older flies, in which niche signals begin to fail as well⁷.

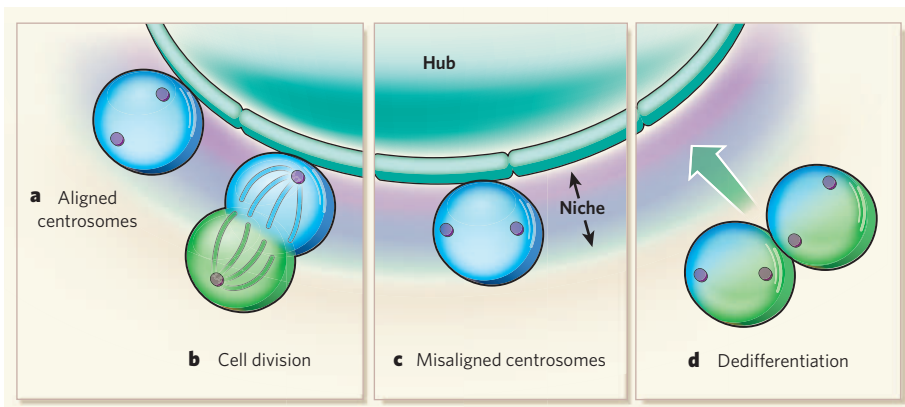


Figure 1 | Centrosome alignment and germline stem-cell (GSC) maintenance in the *Drosophila* testis niche¹. **a**, Normally, GSCs keep one centrosome (purple circle) aligned with the support cells of the hub, so that upon division (**b**) one daughter will remain in the niche while the other will exit and differentiate. **c**, With increasing age, more and more GSCs have misaligned centrosomes — that is, neither is adjacent to the hub — and so do not divide. **d**, Some of these GSCs arise from dedifferentiating older germ cells that re-enter the niche with randomly positioned centrosomes.

A major contribution of this work¹ is the finding that new GSCs are regularly produced by dedifferentiation of downstream germ cells (Fig. 1d). The GSC daughter that leaves the niche divides four times, but incompletely, to generate 2-, 4-, 8- and 16-cell interconnected germline ‘cysts’. It was already known from lineage marking that male GSCs turn over and are replaced at a moderate rate⁸, but the replacement of stem cells by the dedifferentiation of germ cells up to the 8-cell stage had been observed only under experimental conditions^{9,10}.

The data of Cheng *et al.* suggest that, in an old male, all the stem cells that started out in the niche are likely to have undergone replacement by a dedifferentiated cell at least once. Consequently, we now know that GSC activity in male *Drosophila* resides not just in the GSCs themselves, but also in the pool of 2-, 4- and 8-cell germline cysts. This is very similar to a long-standing but controversial model of germ-cell production in the mammalian testes. According to this model, sperm production can be sustained by both individual stem cells and a pool of ‘potential stem cells’ that correspond to the 2-, 4-, 8- and 16-cell germline cysts, which in mice are called A-type spermatogonia^{11,12}. The discovery that dedifferentiation and GSC replacement occur regularly in the normal *Drosophila* testis opens these fascinating processes up to detailed mechanistic study in an accessible system.

How general are these discoveries likely to be? We already know that many other stem cells depend on niches, undergo turnover and replacement, and decline with age¹³. But do they do so in the context of a growing pool of arrested stem cells? Similar cells have not been described during ageing in the *Drosophila* female germ line¹⁴. In the ovary, separate small niches, each supporting just two or three GSCs, are isolated within each of 16 ovarioles. This cellular organization probably reduces the opportunities for contact

between niches and downstream germ cells. In addition, replacement seems to rely on changing the direction of GSC division, to keep both daughters in the niche, rather than on dedifferentiation¹⁵.

But there are many other systems that are compatible with the possibility of such changes. One attractive idea, particularly in large, long-lived animals such as mammals, is that a decline in stem-cell activity with age is actively programmed in order to contain the escalating probability of tumour initiation¹³. It is hard to see how this can be a significant factor in *Drosophila*, however, which is too small and short-lived to suffer significantly from terminal cancer. Short of finding some currently unknown genetic or environmental factors that put Yamashita and colleagues’ flies¹ under stress, we are left with the possibility that stem-cell maintenance is simply less perfect than we might have imagined. Why entrust something as essential as sperm production to a makeshift system that breaks down regularly? “Why indeed?” we hear Jacob asking us. “It works.”

Allan C. Spradling is in the Department of Embryology, Carnegie Institution of Washington, Baltimore, Maryland 21218, USA.
e-mail: spradling@ciwemb.edu

- Cheng, J. *et al.* *Nature* **456**, 599–604 (2008).
- Jacob, F. *Science* **196**, 1161–1166 (1977).
- Yamashita, Y. M., Jones, D. L. & Fuller, M. T. *Science* **301**, 1547–1550 (2003).
- Yamashita, Y. M., Mahowald, A. P., Perlin, J. R. & Fuller, M. T. *Science* **315**, 518–521 (2007).
- Rebollo, E. *et al.* *Dev. Cell* **12**, 467–474 (2007).
- Rusan, N. M. & Peifer, M. *J. Cell Biol.* **177**, 13–20 (2007).
- Boyle, M., Wong, C., Rocha, M. & Jones, D. L. *Cell Stem Cell* **1**, 470–478 (2007).
- Wallenfang, M. R., Nayak, R. & DiNardo, S. *Aging Cell* **5**, 297–304 (2006).
- Kai, T. & Spradling, A. *Nature* **428**, 564–569 (2004).
- Brawley, C. & Matunis, E. *Science* **304**, 1331–1334 (2004).
- Clermont, Y. & Bustos-Obregon, E. *Am. J. Anat.* **122**, 237–247 (1968).
- Nakagawa, T., Nabeshima, Y.-i. & Yoshida, S. *Dev. Cell* **12**, 195–206 (2007).
- Morrison, S. J. & Spradling, A. C. *Cell* **132**, 598–611 (2008).
- Pan, L. *et al.* *Cell Stem Cell* **1**, 458–469 (2007).
- Xie, T. & Spradling, A. C. *Science* **290**, 328–330 (2000).



50 YEARS AGO

During the latter half of October 1957, following the Windscale reactor accident, we observed that the gamma-ray activity of certain members of the staff ... was higher than expected ...

The source of the greater part of this increased activity was identified as iodine-131 ... We decided, therefore, to make some measurements of the iodine-131 in human thyroids ... Eighteen persons were measured, the majority of whom were resident either in, or on the fringe of, the Greater London area ... Since the main route of entry to the human thyroid was via milk, then as the cloud with its iodine-131 from Windscale passed over Britain on October 11, we may expect that the first day on which the milk contained this isotope was October 12 ... If we assume the same pattern of uptake throughout the group, then it is evident that the average person received about 0.04 rad. The magnitude of the radiation dose to adult human thyroids in the people measured was therefore comparable with that expected each year from natural background radiation.

From *Nature* 6 December 1958.

100 YEARS AGO

A Bill for putting in force the decisions of the Berlin Wireless Telegraphy Conference of November, 1906, as embodied in an international convention, has been laid before the French Chamber ... The conference has fixed wave-lengths, one of 300 metres, the other of 600 metres, for the transmission of public messages by the wireless current ... All stations must be able to produce one, at all events, of these two wave-lengths ... Stations on board ship must use the 300-metre wave-length. They are permitted, however, to use other wave-lengths as well, provided that these are under 600 metres. Ships of small tonnage will be allowed to use a wave-length below 300 metres.

From *Nature* 3 December 1908.

50 & 100 YEARS AGO

FERTILIZATION

Welcome to the fold

Paul M. Wassarman

The mammalian egg coat participates in fertilization and prevents more than one sperm from entering the egg. Structural data pinpoint a region common to egg-coat proteins that might mediate these functions.

Eggs are surrounded by one or more extra-cellular coats that take part in fertilization. Mammalian eggs have a single coat, the zona pellucida, which consists of long, interconnected fibrils constructed of a few proteins¹. Mouse zona-pellucida proteins are called ZP1, ZP2 and ZP3, with these last two serving as sperm receptors during fertilization^{1,2}. All three proteins have a characteristic region known as the ZP domain³, which also occurs in related proteins that make up the coat of fish, frog and bird eggs. What's more, the ZP domain is not restricted to egg-coat proteins, but is present in hundreds of extracellular proteins found in virtually all multicellular organisms.

Mutations in the ZP domain can result in severe disorders, such as infertility, deafness and cancer. Nonetheless, no high-resolution structure of either a protein containing the ZP domain or a vertebrate protein involved in fertilization has been obtained. The propensity of ZP-domain proteins to self-assemble into polymers (through the amino-terminal half (ZP-N) of this domain⁴), and to undergo extensive modification after their synthesis from the gene, has made it difficult to solve their structure. But on page 653 of this issue, Monné *et al.*⁵ describe a high-resolution (2.3-ångström) three-dimensional structure of ZP-N from ZP3, revealing that it folds in a novel way.

The ZP domain consists of about 260 amino acids, of which eight are evolutionarily conserved cysteine residues that form disulphide linkages². Proteins containing this domain display a mosaic architecture that is thought to arise from the combination of a common structural building block — the ZP domain — with additional sequences that have specific biological functions.

Monné and colleagues find that two-thirds of ZP-N consists of eight identical β -strands (designated A–E, E', F and G), which are interconnected by loops of varying lengths and shapes. The strands form a sandwich consisting of two sheets, similar to that found in immunoglobulin proteins. However, the sheets rest on a platform created by the extension of strands E', F and G (see Fig. 1 on page 654). Two conserved intramolecular disulphide bonds staple the sandwich together to stabilize the structure. The structural complexity of ZP-N meant that these features could not have been predicted from its amino-acid sequence.

The immunoglobulin superfamily includes proteins that serve as receptors, in cellular

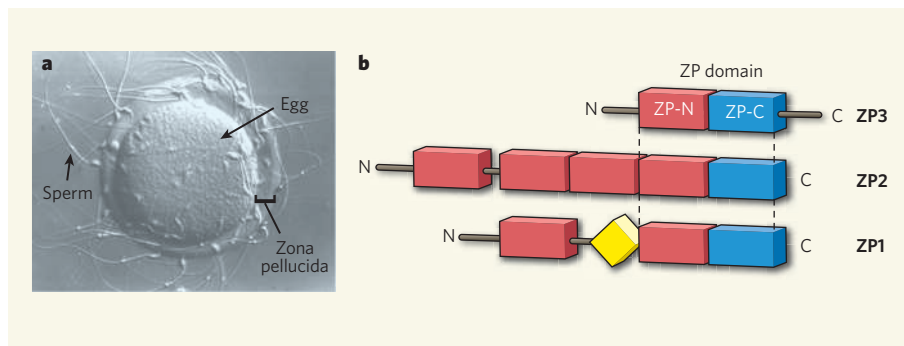


Figure 1 | ZP-domain proteins of the mouse egg. **a**, A mouse egg and sperm cells during fertilization. **b**, The mouse egg zona pellucida consists of three proteins: ZP1, ZP2 and ZP3. All of these proteins contain a ZP domain, which consists of one or several ZP-N regions (red) and a ZP-C region (blue). In addition, ZP1 has a domain called the trefoil domain (yellow). Monné *et al.*⁵ report the three-dimensional structure of the ZP-N region of ZP3.

adhesion and other processes. Although ZP-N adopts an immunoglobulin-like fold, overall it has little structural similarity to other immunoglobulin-like domains and represents a new domain within this superfamily of proteins⁶. Nonetheless, Monné and colleagues' data⁵ indicate that the platform region of the molecule participates in intermolecular interactions. This region contains two conserved tyrosine residues that interact with each other. When one of the two tyrosines is mutated⁷, or when either of them is missing⁸, ZP-domain proteins fail to form polymers. Collectively, these observations strongly suggest that the unique fold in ZP-N triggers the polymerization of ZP-domain proteins into higher-order structures. In the zona pellucida, for instance, this region might mediate the polymerization of the ZP2 and ZP3 dimers that form the long fibrils interconnected by ZP1 (ref. 2).

Monné *et al.* also present a structural model that takes into account the proposal⁹ that egg-coat proteins such as ZP1 and ZP2 in frogs, birds and mammals have several copies of ZP-N upstream of the ZP domain (Fig. 1). Only ZP-domain proteins involved in fertilization contain the extra copies of ZP-N, and in ZP2, these form extended, rod-like structures. It is possible that interaction between these structures on different ZP2 molecules accounts for the post-fertilization compaction of zona-pellucida fibrils that creates a barrier to further sperm penetration and masks sperm-binding sites on ZP-N. The copies of ZP-N have an unusually high degree of amino-acid diversity among different species. Amino-acid variation also occurs in egg-coat proteins that mediate

sperm recognition in both vertebrates and invertebrates, and this variation is thought to be important for speciation^{10,11}. So perhaps amino-acid diversity in the ZP-N repeats of ZP2 are necessary for species-restricted fertilization in mammals.

Monné and colleagues' observations provide intriguing insights into the polymerization of ZP-domain proteins in general and aspects of mammalian fertilization in particular. A full understanding of the relationships between ZP2 and ZP3 in zona-pellucida fibrils, however, will require further structural studies of the entire ZP domain. For example, it is unclear why both ZP2 and ZP3 are required for the zona pellucida to assemble around the egg²; certainly, some ZP-domain-containing proteins, such as uromodulin, can polymerize into fibrils and form a matrix on their own. Perhaps the carboxy-terminal region of the ZP domain of ZP2 and ZP3 imparts specificity to interactions between the proteins. The structure of an entire ZP domain could shed light on this and other issues.

Determining high-resolution structures of ZP2 and ZP3 should be most illuminating. In the case of ZP3, the polypeptide downstream of the ZP domain serves as an initial binding site for sperm, whereas the polypeptide in the ZP domain of ZP2 — including the extra copies of ZP-N — provides subsequent binding sites for sperm. So stay tuned for the exciting instalments that are sure to follow this tantalizing first glimpse of zona-pellucida protein structures⁵.

Paul M. Wassarman is in the Department of Developmental and Regenerative Biology, Mount Sinai School of Medicine, 1 Gustave L. Levy

Place, New York, New York 10029, USA.

e-mail: paul.wassarman@mssm.edu

1. Florman, H. M. & Ducibella, T. in *Physiology of Reproduction* Vol. 1, 2nd edn, 55–112 (Academic, 2006).
2. Wassarman, P. M. *J. Biol. Chem.* **283**, 24285–24289 (2008).
3. Jovine, L. *et al.* *Annu. Rev. Biochem.* **74**, 83–114 (2005).
4. Jovine, L. *et al.* *BMC Biochem.* **7**, 11 (2006).
5. Monné, M., Han, L., Schwend, T., Burendahl, S. & Jovine, L. *Nature* **456**, 653–657 (2008).

6. Barclay, A. *Semin. Immunol.* **15**, 215–223 (2003).
7. Legan, P. K. *et al.* *Nature Neurosci.* **8**, 1035–1042 (2005).
8. Ten Dijke, P., Goumans, M. J. & Pardali, E. *Angiogenesis* **11**, 79–82 (2008).
9. Callebaut, I., Mornon, J. P. & Monget, P. *Bioinformatics* **23**, 1871–1874 (2007).
10. Swanson, W. J. *et al.* *Proc. Natl Acad. Sci. USA* **98**, 4639–4643 (2001).
11. Galindo, B. E., Vacquier, V. D. & Swanson, W. J. *Proc. Natl Acad. Sci. USA* **100**, 4639–4643 (2003).

only a brilliant idea for many decades because of the technical challenges involved in observing such low-surface-brightness echoes. Now, after 40 years, and thanks to the enormous technological progress and the larger dimensions of modern telescopes, Shklovski and van den Bergh's prediction has finally proved to be correct^{9,10}.

But what is a supernova light echo? A supernova produces a flash of optical radiation that propagates in all directions. If the blast occurred in a vacuum, the observer would detect the event through photons that travel unperturbed along the line of sight. But interstellar space is not empty, and thus photons propagating away from the line of sight are likely to encounter dust clouds that scatter their light in all directions. Therefore, a small fraction of the incoming radiation is expected to be seen by the observer as a faint, local re-brightening in the proximity of the explosion site. Because of the finite speed of light, scattered photons will reach Earth after some delay compared with those travelling along a straight, unperturbed path (Fig. 1). Because this phenomenon is similar to that well known for acoustic waves, it is labelled a light echo.

Interestingly, as the flash propagates through the cloud, the re-brightening shifts in space (Fig. 1), giving the impression that material is moving at velocities that can actually exceed the speed of light. In reality, the effect is purely geometrical and nothing physically moves,

ASTROPHYSICS

Echo from an ancient supernova

Andrea Pastorello and Ferdinando Patat

Light reflected off a dust cloud in the vicinity of the relic of Tycho Brahe's supernova, whose light first swept past Earth more than four centuries ago, literally sheds light on the nature of this cosmic explosion.

The occurrence of a supernova in our Galaxy is a rare event — only six such supernovae were registered during the past millennium¹. Although their remnants have been extensively studied, the nature of the original explosions is the subject of debate. Supernovae are classified into different types mainly on the basis of their observed spectral properties, and occasionally through the analysis of well-sampled light curves (plots that show the evolution of their brightness over a period of time). The classification of historical supernovae — those whose light was first detected on Earth centuries ago — is more problematic than that of events detected recently, because the sparse visual observations reported in historical records are often difficult to calibrate, and the resulting light curves are rather uncertain. In addition, the study of their relics does not always give unequivocal answers as to which category they fall into.

On page 617 of this issue, Krause *et al.*² offer convincing evidence that the supernova first observed by Tycho Brahe in 1572 (SN1572) was a type Ia supernova — one originated by the thermonuclear explosion of a white dwarf star in a binary system. The authors' investigation was triggered by a recent paper³ that announced the discovery of light echoes from two ancient supernovae, one of which was SN1572. The luminous echo from Tycho Brahe's supernova was visible in recent images of its remnant as an extended light source which, according to astrometric observations that determine the precise positions of the source in the sky, seemed to shift with time through a dust cloud in the remnant's vicinity^{2,3}. Krause and colleagues classified SN1572 through a spectrum of its echo, a method with proven credentials in classifying ancient supernovae. For example, the youngest supernova remnant in our Galaxy, Cassiopeia A (ref. 4), was classified in this way as a type IIb (ref. 5) event (a supernova generated by the gravitational collapse of the

core of a massive, moderately hydrogen-poor star), and the relic SNR0509–675 in the Large Magellanic Cloud was classified as a luminous type Ia supernova⁶.

During the 1960s, Iosif Shklovski⁷ and Sidney van den Bergh⁸ independently showed that the detection and subsequent follow-up spectroscopic analysis of scattered-light echoes in the surroundings of supernova sites could be used to get information on the explosions that generated them. But this technique remained

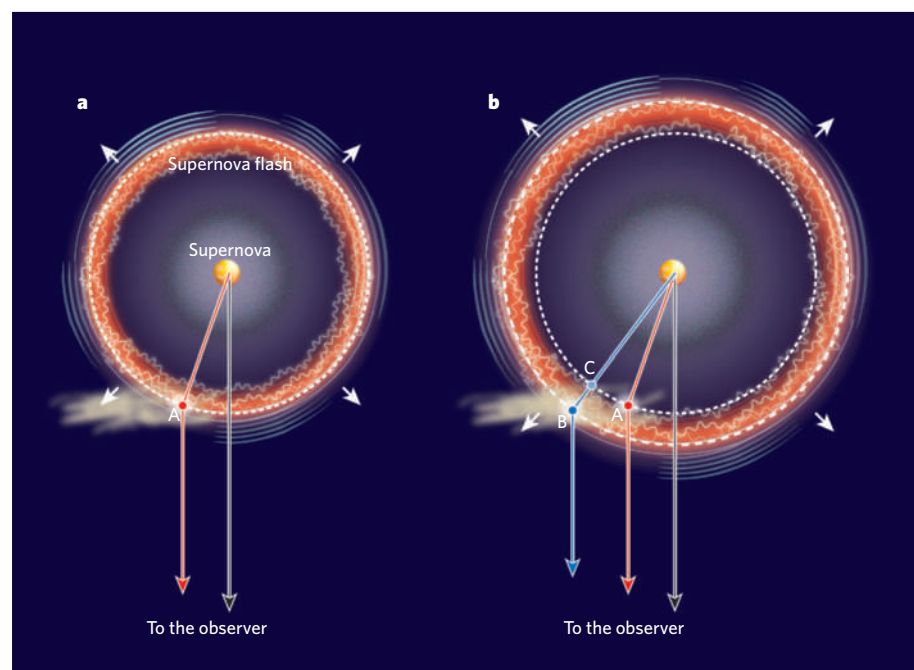


Figure 1 | Catching the echoes. As a supernova flash expands at the speed of light, it scatters off and illuminates different regions of a nearby dust cloud, creating delayed re-brightenings, the 'echoes', across the cloud. **a**, A light echo (red) reaches Earth with a given delay with respect to the original light that travels along a straight, unperturbed path (black). **b**, A subsequent light echo is observed with a further delay (blue). The material in the cloud seems to move from A to B (but it does not) at a super-luminal speed because the apparent distance, as measured by the observer, that the light travelled (the path from A to B) is larger than the real distance spanned by photons (from C to B). The spectral analysis of a light echo from Tycho Brahe's supernova performed by Krause *et al.*² provides information about the nature of this supernova centuries after it was first detected on Earth.

but variations of brightness with time allow astronomers to detect light echoes through reiterated observations of selected areas in the sky. A time-evolving re-brightening is exactly what Krause *et al.*² observed in a particular (and limited) region of the remnant of SN1572.

The light echo thus carries a sort of fossil imprint of the original supernova, and so analysis of the reflected light can unveil characteristics of it even centuries after it was first detected on Earth. Krause *et al.* obtained a spectrum of the bright-light echo of SN1572 with the Subaru 8.2-metre telescope on Mauna Kea (Hawaii) and showed that prominent silicon, sulphur and iron lines — which are typical of a normal type Ia supernova around its maximum — were present in the spectrum.

Type Ia supernovae are thought to be generated by the explosion of carbon–oxygen white dwarf stars in binary systems. As the white dwarf accretes mass from its companion star and eventually reaches the critical Chandrasekhar mass (1.4 solar masses), a thermonuclear explosion disrupts the star, releasing materials that are the product of the stellar nuclear fusion (mostly silicon and elements of the iron group). Nickel (⁵⁶Ni) in particular is ejected in great amounts — up to 50% of the whole white dwarf's mass. This radioactive isotope decays initially into radioactive cobalt (⁵⁶Co), and then into stable iron (⁵⁶Fe). Such a decay chain provides enough energy to keep the material ejected by the explosion hot for several months, and supports the supernova luminosity in the early phase of the transition towards the remnant stage. During this transformation, some — but not all — information about the explosion itself is lost. This implies that, in principle, the type of original supernova cannot be uncovered by studying the chemical abundances in the remnant alone.

Nevertheless, Tycho Brahe's supernova was suspected to be of type Ia well before the observations reported by Krause *et al.*². The historical light curve¹¹, the studies of the remnant in the radio¹² and X-ray¹³ wavelengths, and the discovery of a surviving candidate G-type companion star¹⁴ (a star just like our Sun) all suggested a thermonuclear supernova rather than a core-collapse event. The fundamental contribution of Krause and colleagues' work is to transform these clues into definite proof — we are now fully confident that one of the most popular supernova remnants detected in our Galaxy was produced by an ordinary type Ia supernova that was first detected more than 400 years ago.

The technique of observing light echoes from supernovae is a remarkable observational tool with which to pigeonhole the type of supernova. It will allow astrophysicists to characterize other supernova remnants in our Galaxy and in nearby galaxies. This will hopefully clarify the relationship between supernova relics and their explosion mechanisms. Finally, it is likely that precise information about the

frequency of the different supernova types in our Galaxy and its surroundings will shed light on the star-formation history and chemical evolution of the Local Group of galaxies. ■

Andrea Pastorello is at the Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK. Ferdinando Patat is at the European Organization for Astronomical Research in the Southern Hemisphere (ESO), Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany.

e-mails: a.pastorello@qub.ac.uk;
fpatat@eso.org

- Green, D. A. & Stephenson, F. R. in *Lecture Notes in Physics* Vol. 598 (ed. Weiler, K.) 7–19 (Springer, 2003).
- Krause, O. *et al. Nature* **456**, 617–619 (2008).
- Rest, A. *et al. Astrophys. J.* **680**, 1137–1148 (2008).
- Thorstensen, J. R., Fesen, R. A. & van den Bergh, S. *Astron. J.* **122**, 297–307 (2001).
- Krause, O. *et al. Science* **320**, 1195–1197 (2008).
- Rest, A. *et al. Astrophys. J. Lett.* **681**, L81–L84 (2008).
- Shklovski, I. S. *Astron. Circ. (USSR)* **306**, 2–4 (1964).
- van den Bergh, S. *Publ. Astron. Soc. Pacif.* **77**, 269–271 (1965).
- Rest, A. *et al. Nature* **438**, 1132–1134 (2005).
- Krause, O. *et al. Science* **308**, 1604–1606 (2005).
- Baade, W. B. *Astrophys. J.* **102**, 309–317 (1945).
- Baldwin, J. E. & Edge, D. O. *Observatory* **77**, 139–143 (1957).
- Hughes, J. P. *et al. Astrophys. J.* **444**, L81–L84 (1995).
- Ruiz-Lapuente, P. *et al. Nature* **431**, 1069–1072 (2004).

BEHAVIOURAL ECOLOGY

The social side of wild yeast

David C. Queller

The workhorse of cell biology, yeast, is a surprisingly cooperative organism. It uses an unusual means of identifying partners — a 'green-beard gene', which encodes a tag that must match among cooperating cells.

Everyone knows how a glass or two of beer can act as social glue, making even misanthropes amiable. Oddly, the production of beer has a similarly convivial effect on the tiny brewers that make it, cells of the yeast *Saccharomyces cerevisiae*. As alcohol content rises, the normally solitary cells begin to adhere to each other in clumps called flocs. Work on these flocs, just published by Smukalla *et al.* in *Cell*¹, shows that the yeast cells face a familiar social dilemma, but that they solve it in an exotic fashion.

In addition to its humble jobs as brewer and baker, yeast has added a high-tech career. It has become the principal laboratory model organism for studying the biology of eukaryotes — organisms such as plants and animals that have a membrane-bound nucleus. We have therefore come to understand yeast as well as any organism. Yet, although brewers have known about flocculation for centuries — sedimentation of the flocs provides an easy way to remove the yeast and keep the beer from tasting like the yeast paste Marmite — biologists have been slower off the mark. The problem is that flocculation was lost during the domestication of yeast. Smukalla *et al.*¹ therefore chose to study feral strains. They show that expression of one of five flocculation cell-adhesion genes, *FLO1*, explains much of the variation in flocculation (Fig. 1, overleaf). They establish causation by expressing *FLO1* in the domesticated strain, resurrecting flocculation, and putting *FLO1* into *Saccharomyces paradoxus*, inducing flocs even in this non-flocculating species.

Flocculation is a true cooperative trait that poses the classical social dilemma of how to sustain cooperation in the face of cheaters.

Smukalla *et al.* show that cells on the inside of flocs are protected against damage from chemicals, including alcohol, partly because of physical shielding provided by cells on the outside. But those gains come at a cost; cells that express *FLO1* grow more slowly, even if they are prevented from flocculating. So, is it possible for cheaters to gain the protection of flocs without paying the cost?

The answer seems to be no, according to Smukalla *et al.*, because the ticket for admission to the floc is the ability to make the adhesion protein. It is just as if the bonhomie of the beer hall extended only to those drinking the same brand of beer or, on St Patrick's Day, only to those drinking green beer. In behavioural ecology, such tokens of inclusion are named not for green beer but for green beards. The name is fanciful because, until recently, such genes were purely imaginary. It is well known that genes for helping can spread by benefiting relatives, who share the gene with a specifiable probability. The alternative is to identify and help actual bearers of the gene, whether they are relatives or not². Richard Dawkins³ argued that such green-beard genes were unlikely to occur because they would need to cause three traits: a label, such as a green beard; recognition of others with the label; and nicer behaviour towards those with the label.

Green-beard systems have, however, begun to leave the realm of the imaginary. The initial examples involved nasty behaviour towards those lacking the gene. In the fire ant *Solenopsis invicta*, a linked set of alleles, including one encoding an olfactory receptor, causes workers to kill queens that lack the green-beard allele⁴. Microbes are proving a richer source of

green-beard genes. Many bacteria act rather like fire ants, using linked poison-antidote genes to kill members of the same species that don't possess the antidote gene⁵. The social amoeba *Dictyo steliundiscodeum* provided the first single-gene example and the first altruistic one. Its *csaA* gene codes for a homophilic adhesion protein, one that grabs on to the same protein on other cells, so excluding strains that do not express the protein from the benefits of later altruism within the group⁶.

The yeast *FLO1* green-beard gene adds a new level of interest because it is highly variable, and because the variation — in the number of repeats of a 100-base sequence — exerts a major influence on protein-binding strength. Conceivably, such variation could lead to multiple recognition tags (green beards, red beards, yellow beards and so on), although this possibility remains to be tested. Alternatively, the variation might be the product of highly complex adaptive dynamics, which have been observed in simulation models as various green-beard cooperators and cheaters rise and fall in frequency⁷.

The discovery of *FLO1* also extends the reach of single-gene green beards. It is easy to see how homophilic adhesion molecules such as the product of *csaA* could recognize each other. In fact, it was predicted that they could be green beards⁸. But *FLO1* makes a heterophilic adhesion protein, one that binds not to itself but to oligosaccharides on the walls of other cells. Cells without the adhesion protein

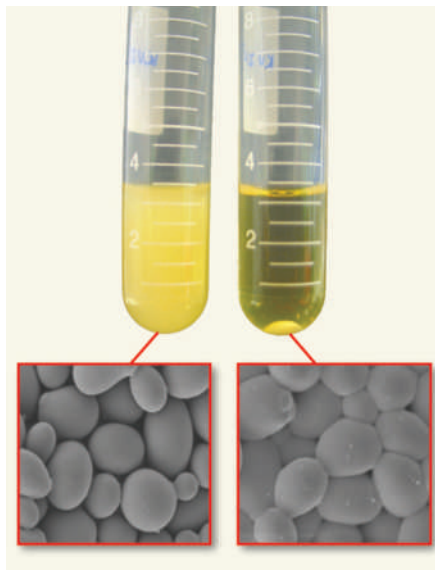


Figure 1 | Yeast of a feather flock together. In the yeast *Saccharomyces cerevisiae*, the *FLO1* gene, whose expression has been lost in the domesticated strain (left), causes cells to adhere together in flocs that form a protective sediment (right). As Smukalla *et al.* report¹, *FLO1* is a green-beard gene that benefits copies of itself rather than kin as such.

can still be bound by those that have it, but the binding is weaker than proper two-way binding, so such cells usually remain outside flocs. Even worse, the flocless cells that do get into flocs are exploited — they tend to end up in

the outer layer, where they can be damaged in protecting the floc cells inside.

The discovery of a green beard in an organism as mundane as yeast suggests that such genes might prove to be quite common, at least in microbes. More generally, microbial sociality itself is probably more common than lab studies might indicate. In model systems, domestication can easily lead to loss of social behaviours⁹. Another example is the bacterium *Bacillus subtilis*, in which multicellular fruiting bodies were seen only when wild strains were studied¹⁰. Much like the entrepreneur who rescued Marmite from the discards of the brewery, scientists can find value in the traits discarded by the domesticators.

David C. Queller is in the Department of Ecology and Evolution, MS-170, Rice University, Houston, Texas 77251-1892, USA.

e-mail: queller@rice.edu

1. Smukalla, S. *et al.* *Cell* **135**, 726–737 (2008).
2. Hamilton, W. D. J. *Theor. Biol.* **7**, 17–52 (1964).
3. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, 1976).
4. Krieger, M. J. B. & Ross, K. G. *Science* **295**, 328–332 (2002).
5. Riley, M. A. & Wertz, J. E. *Annu. Rev. Microbiol.* **56**, 117–137 (2002).
6. Queller, D. C. *et al.* *Science* **299**, 105–106 (2003).
7. Jansen, V. A. A. & van Baalen, M. *Nature* **440**, 663–666 (2006).
8. Haig, D. *Proc. Natl Acad. Sci. USA* **93**, 6547–6551 (1996).
9. Velicer, G. J., Kroos, L. & Lenski, R. E. *Proc. Natl Acad. Sci. USA* **95**, 12376–12380 (1998).
10. Branda, S. S. *et al.* *Proc. Natl Acad. Sci. USA* **98**, 11621–11626 (2001).

NEUROSCIENCE

Along memory lane

Yukiko Goda

Memories are encoded by efficient signalling between neurons. The myosin V proteins help this process by shuttling receptors and membranes to make synaptic junctions better detectors of incoming signals.

Synaptic junctions transmit information between neurons. The efficiency with which they do this is affected by how frequently they are activated, a cellular equivalent of experience. For example, repeated activation yields a lasting increase in the efficiency of synaptic transmission — a process called long-term potentiation (LTP) — which is thought to underlie memory formation. LTP depends both on enhanced insertion of receptors for the neurotransmitter glutamate at spines (the sites of synapses) and on spine growth. How neurons coordinate these two processes to trigger LTP has remained a mystery. Writing in *Cell*, Wang, Ehlers and colleagues¹ demonstrate that the motor protein myosin Vb is one missing link.

Dedicated to receiving signals, spines are tiny protrusions on the dendritic processes of

neurons. They contain two types of glutamate receptor: AMPA receptors, which mediate most glutamate-dependent synaptic transmission; and NMDA receptors. On induction of LTP, activated NMDA receptors mediate a rapid influx of calcium ions (Ca^{2+}) into spines, which initiates a cascade of signalling events, culminating in the insertion of extra AMPA receptors into the spines, and spine growth.

Previously, Ehlers and colleagues had shown^{2,3} that, during LTP, organelles known as recycling endosomes, which are normally found in dendritic regions near the base of spines, provide both AMPA receptors and membrane material for spine growth. In their latest work, the authors asked¹ what molecule directs the transport of recycling endosomes from dendrites into spines in response to Ca^{2+} signals.

Within the cell, motor proteins move organelles along tracks of cytoskeletal structures called microtubules and actin filaments. Dendrites are rich in microtubules, whereas spines are mostly filled with actin. So the ideal candidate for Wang and colleagues' search would be a Ca^{2+} -dependent protein that could hijack recycling endosomes from microtubules in dendrites and move them along actin within spines.

Myosin V proteins seem to fit the bill. In many cell types, these actin-based motor proteins specialize in organelle transport, their structures are regulated by Ca^{2+} , and they can counteract organelle movements along microtubules⁴. Furthermore, one of the three myosin V isoforms, myosin Vb, binds to recycling endosomes through the Rab11 protein on these organelles and another protein known as Rab11-FIP2, which interacts with Rab11 (ref. 4). In neurons, myosin-Vb activity affects surface expression and clustering of AMPA receptors⁵.

Wang *et al.* find that myosin Vb is indeed the protein mediating AMPA-receptor and membrane trafficking to spines. Visualizing the cellular movement of this protein by tagging it with a fluorescent molecule, they show that, under resting conditions, it is mostly present in spines. However, on induction of

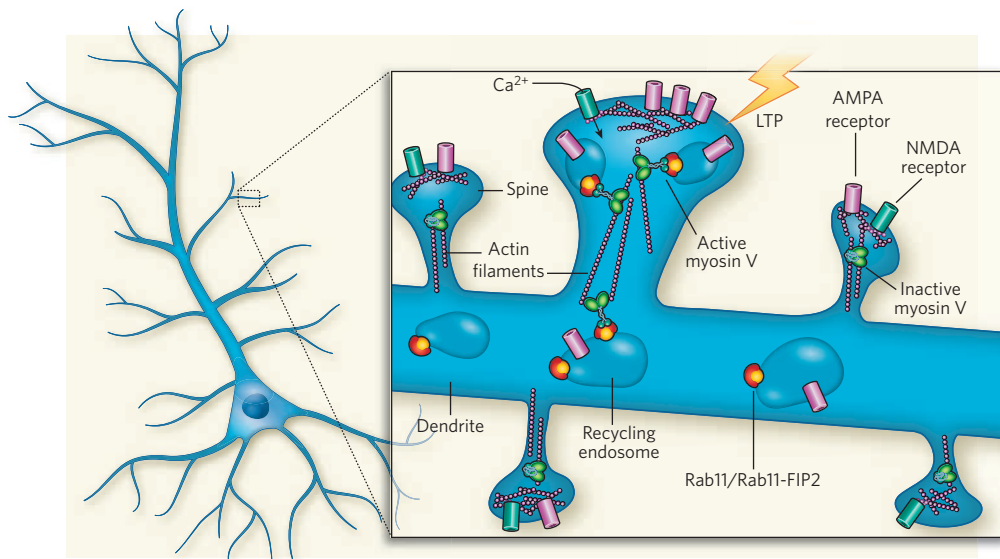


Figure 1 | Molecular basis of long-term potentiation (LTP). Neurons receive synaptic transmission from other neurons (not shown) at dendritic spines. Strengthening of synapses during LTP — a cellular correlate of memory — requires the delivery of receptors and membrane pieces to spines by a myosin V motor protein^{1,7}. Influx of Ca^{2+} through NMDA receptors activates myosin V in spines. The active motor protein then moves to the dendrite shaft and binds to Rab11/Rab11-FIP2 on recycling endosomes containing AMPA receptors. Finally, it transports the cargo into and along spines via actin filaments to mediate insertion of AMPA receptors at the cell surface, as well as spine growth through membrane insertion.

LTP and activation of NMDA receptors, it moves to dendrites, binds to recycling endosomes through Rab11/Rab11-FIP2 and then returns to spines with its cargo containing the AMPA-receptor subunit GluR1 (Fig. 1).

Myosin Vb can adopt two conformations: a folded, inactive state with a high affinity for actin, and an extended state, which binds to Rab11/Rab11-FIP2. The LTP-triggered increase in Ca^{2+} concentrations seems to affect myosin Vb by inducing a conformational change in this protein from the inactive to the extended state, allowing it to connect to recycling endosomes¹.

But is myosin-Vb activity required for induction of LTP? To propel their cargo, myosin V proteins use ATP molecules as fuel. The ATP-binding pocket of these proteins can be subtly mutated to preserve their normal function, but to allow the protein to be quickly locked by chemical means into the inactive actin-bound form⁶. This approach avoids potential complications due to compensatory mechanisms that could arise during long-term manipulations of myosin V activity. Using engineered mice with such mutations, Wang *et al.*¹ measured synaptic efficiency in mouse brain slices. They show that acute inhibition of myosin Vb blocks LTP within minutes.

Another recent study⁷ has also demonstrated a role for myosin Va — another myosin-V isoform — in LTP. With the help of active Rab11, this protein binds directly to the cytoplasmic tail of GluR1, supporting the connection between myosin isoforms, AMPA receptors and recycling endosomes. Nevertheless, brain slices of a mouse mutant lacking myosin Va show normal LTP (ref. 8), suggesting that the role of myosin Va in LTP is subtle.

It would be interesting to test whether, using Wang and colleagues' approach, acute inhibition of myosin Va over a timescale of minutes would be sufficient to block LTP.

A key feature of LTP is its input specificity — that is, an increase in synaptic strength is confined to spines in which LTP has been induced. Because spines are present at high density, how does myosin Vb deliver its cargo to where it is needed? The protein could use local Ca^{2+} gradients, specific to active spines, to adjust and direct its motility. But LTP is not alone in its requirement for Ca^{2+} influx through NMDA receptors. Another form of lasting change called long-term depression (LTD), which weakens synaptic strength by removing AMPA receptors, also requires NMDA-receptor activation and is input specific. So myosin Vb may be tuned to preferentially detect LTP-associated Ca^{2+} dynamics in specific spines and/or local dendrites, and work in concert with other signals associated only with LTP or LTD. For instance, spines undergoing LTD could generate signals that block the entry of recycling endosomes. A detailed look at how Ca^{2+} dynamics at specific spines controls the local distribution of myosin V and its GluR1-containing cargo would be valuable.

Identification of an actin motor protein as the delivery vehicle for membranes and AMPA receptors to spines during LTP highlights the significance of actin polymerization in spine growth. Actin is highly enriched in spines, in which at least two distinct pools of it are found: stable actin filaments within the spine core and a dynamic pool at the spine 'head', where it could organize glutamate receptors by interacting with other scaffold proteins⁹. LTP-associated spine growth — often an enlargement of

the spine head — probably involves polymerizing a new set of actin filaments⁹.

How does myosin Vb cope with different pools of actin? When entering spines, this protein could carry recycling endosomes along the stable actin at the core. On reaching the head region, actin polymerization itself could help to propel the recycling endosomes (or part of them) towards the spine periphery by elongating filaments behind the cargo. This would ensure that membrane insertion is synchronized with the growth of the actin cytoskeleton, thus preventing the inserted membranes from collapsing. Alternatively, a distinct motor protein or trafficking machine could catalyse the final cell-surface delivery of membranes, with additional mechanisms in place for coupling membrane insertion and actin polymerization.

The abundance of GluR1 at the spine surface must also be tightly controlled with respect to spine growth^{10,11}. Curiously, Wang *et al.*¹ find that, on induction of LTP, the extent of GluR1 insertion and that of bulk endosomal membrane do not always go hand in hand. Compared with bulk endosomal traffic, therefore, GluR1 may be sorted into a subpopulation of recycling endosomes that take a different course to reach the cell surface. Alternatively, the extent of GluR1 accumulation on the cell surface could be restricted.

Other questions also remain. How do myosin V proteins select their cargo? What are the specific functions of their three isoforms in AMPA-receptor and membrane trafficking, and are their activities coupled to the membrane-fusion machinery? In addition, myosin V may have other LTP-related functions at spines. For example, that myosin Va is involved in localizing messenger RNA to spines¹² suggests its involvement in translation-dependent forms of LTP. It is likely that synapses will remain a gold mine for revealing the molecular repertoire of memory mechanisms for years to come. ■

Yukiko Goda is in the MRC Laboratory for Molecular Cell Biology and Cell Biology Unit, Department of Neuroscience, Physiology and Pharmacology, University College London, Gower Street, London WC1E 6BT, UK. e-mail: y.goda@ucl.ac.uk

1. Wang, Z. *et al.* *Cell* **135**, 535–548 (2008).
2. Park, M., Penick, E. C., Edwards, J. G., Kauer, J. A. & Ehlers, M. D. *Science* **305**, 1972–1975 (2004).
3. Park, M. *et al.* *Neuron* **52**, 817–830 (2006).
4. Desnos, C., Huet, S. & Darchen, F. *Biol. Cell* **99**, 411–423 (2007).
5. Lisé, M.-F. *et al.* *J. Biol. Chem.* **281**, 3669–3678 (2006).
6. Provance, D. W. *et al.* *Proc. Natl Acad. Sci. USA* **101**, 1868–1873 (2004).
7. Correia, S. S. *et al.* *Nature Neurosci.* **11**, 457–466 (2008).
8. Schnell, E. & Nicoll, R. A. *J. Neurophysiol.* **85**, 1498–1501 (2001).
9. Cingolani, L. A. & Goda, Y. *Nature Rev. Neurosci.* **9**, 344–356 (2008).
10. Kopec, C. D., Li, B., Wei, W., Boehm, J. & Malinow, R. *J. Neurosci.* **26**, 2000–2009 (2006).
11. Kopec, C. D., Real, E., Kessels, H. W. & Malinow, R. *J. Neurosci.* **27**, 13706–13718 (2007).
12. Yoshimura, A. *et al.* *Curr. Biol.* **16**, 2345–2351 (2006).

Pleiotropic scaling and QTL data

Arising from: G. P. Wagner *et al.* *Nature* **452**, 470–473 (2008)

Wagner *et al.*¹ have recently introduced much-needed data to the debate on how complexity of the genotype–phenotype map affects the distribution of mutational effects. They used quantitative trait loci (QTLs) mapping analysis of 70 skeletal characters in mice² and regressed the total QTL effect on the number of traits affected (level of pleiotropy). From their results they suggest that mutations with higher pleiotropy have a larger effect, on average, on each of the affected traits—a surprising finding that contradicts previous models^{3–7}. We argue that the possibility of some QTL regions containing multiple mutations, which was not considered by the authors, introduces a bias that can explain the discrepancy between one of the previously suggested models and the new data.

Wagner *et al.*¹ define the total QTL effect as

$$T = \sqrt{\sum_{i=1}^N A_i^2}$$

where N measures pleiotropy (number of traits) and A_i the normalized effect on the i th trait. They compare an empirical regression of T on N with predictions from two models for pleiotropic scaling. The Euclidean superposition model^{3,4} assumes that the effect of a mutation on a given trait is independent of its effect on other traits. It predicts independence of A_i and N and therefore a square-root scaling of T with N . The total effect model^{5–7} assumes that the total effect T is independent of the number of affected traits N . As their data show a roughly linear dependence of T on N , the authors reject both models. They conclude that not only the total effect, but, surprisingly, even the effect per trait increases with pleiotropy.

Although the authors address several problems in interpreting QTL data, there is one confounding factor not considered that could explain the most surprising part of their results. Crucially, whereas models address pleiotropy of single mutations, the data presented by Wagner *et al.*¹ are on QTL pleiotropy. As QTLs correspond to relatively large genomic regions, they can contain more than a single mutation. The pleiotropic QTLs in this study correspond to regions of 3–50 centimorgans (ref. 2). Major QTLs can break down into the contributions of several genes⁸ and even multiple mutations in a single gene⁹, sometimes affecting different traits¹⁰. Multiple-mutation QTLs are likely in the present study^{1,2}, for several reasons: first, the data are based on F2 mapping, which has limited power to resolve the effects of linked loci². Second, as pleiotropy is the null-model², there is a bias for grouping several QTLs into a single one. There is evidence that this may indeed be the case for all highly pleiotropic QTLs². Third, for a single trait, a small number of QTLs is the null-model and indeed the study does not allow for more than two QTLs per trait on the same chromosome². Last, fixation of different mutations at the same gene in the high- and low-selection lines seems plausible.

How would multiple mutations in some QTLs affect the results of the paper? The finding of low levels of pleiotropy is conservative. Indeed, pleiotropy of the single mutations can only be lower than pleiotropy of the QTLs. The effect on the scaling behaviour is more involved. Even if the effects of single mutations (and single-mutation QTLs) strictly follow the square-root scaling predicted by the superposition model, this is generally not true for QTLs with several mutations (Fig. 1). The predicted square-root dependency is reproduced only if the sets of traits affected by the different mutations in a QTL are non-overlapping. Overlapping sets of traits may produce an almost linear relationship of N and T , and thus an apparent deviation from the superposition model (Fig. 1).

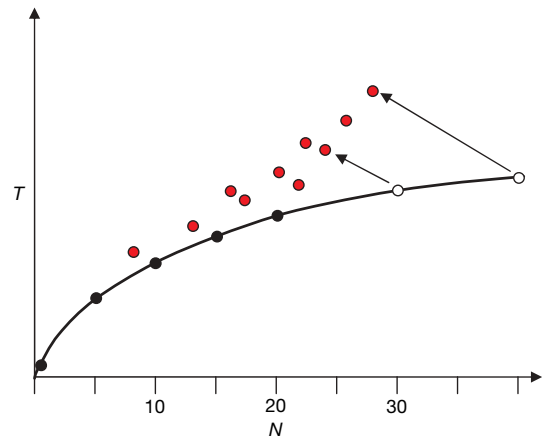


Figure 1 | Impact of multiple mutations on the total QTL effect, T . For all mutations, strict scaling according to the Euclidean superposition model is assumed. Black filled circles, single-mutation QTLs. Open circles, QTLs with two mutations that affect non-overlapping traits. Red filled circles, total effects of double-mutant QTL with overlapping trait ranges (see Methods). Arrows, shift due to the overlap.

Overlaps among sets of traits affected by different mutations in a single QTL are likely, particularly if each mutation affects many traits. Vice versa, QTLs with multiple mutations will (on average) affect a larger number of traits. We therefore expect the largest deviations from an overall square-root relation for data points with large N . This possibly explains the large deviations for the four QTLs with $N > 25$ of Wagner *et al.*¹. For example, a two-mutation QTL where both mutations affect ~ 20 traits could reproduce the observed results.

The T – N regression is less affected by multiple-mutation QTLs if T is measured in the linear Manhattan metric¹, $T = \sum_i |A_i|$. Here, the superposition model predicts a linear regression. Overlaps in the trait sets of multiple-mutation QTLs lead to reduced N , but do not affect T . Indeed, the deviation of the data from a linear regression line appears to be small¹.

Could multiple-mutation QTLs even reconcile the total effect model with the observed data? The answer is, in principle, yes, but only under extreme assumptions. In fact, we would have to explain the entire increase in T with N as solely due to ever-larger numbers of mutations in more pleiotropic QTLs.

Thus the data¹ provide evidence against the total effect model, but not against the Euclidean superposition model. It therefore seems premature to accept a linear regression of the total mutational effect on the degree of pleiotropy or to ponder why mutations with higher pleiotropy should have a larger effect on each affected trait. Eventually, data from QTL fine-scale mapping will help resolve these issues by determining the number of mutations per QTL and the correlation among sets of traits affected by mutations in the same QTL region.

METHODS

Assume that all trait-specific mutational effects are equal to some constant a . The total effect t of a mutation affecting n traits is:

$$t = a\sqrt{n}$$

For single-mutation QTLs, $N = n$ and $T = t$, and data points follow the square-root relation predicted by the superposition model (Fig. 1, black filled circles). The total effect T of a multiple-mutation QTL $\{n_1, n_2, \dots\}$ follows the same

square-root law only if $N = \sum_i n_i$, that is, trait sets affected by the different mutations are non-overlapping (open circles in Fig. 1 for QTL {20,10}, {20,20}).

For overlapping trait sets, $N < \sum_i n_i$ and $T > a\sqrt{\sum_i n_i}$. For a two-mutation QTL with overlap n_{12} , $N = n_1 + n_2 - n_{12}$ and $T = a\sqrt{n_1 + n_2 + 2n_{12}}$. Red filled circles in Fig. 1 show T and N of all two-mutation QTLs with $n_1 \geq n_2 \in \{5, 10, 15, 20\}$. The assumed overlap is $n_{12} = p(n_1) \cdot n_2$ with $p(n_1) = 0.6, 0.5, 0.4, 0.3$ for $n_1 = 20, 15, 10, 5$, respectively.

Joachim Hermisson¹ & Alistair P. McGregor²

¹Fakultät für Mathematik and Max F. Perutz Laboratories, University of Vienna, Nordbergstr. 15, 1090 Vienna, Austria.

e-mail: joachim.hermisson@univie.ac.at

²Institut für Populationsgenetik, Veterinärmedizinische Universität Wien, Veterinärpl. 1, 1210 Vienna, Austria.

Received 29 May; accepted 2 September 2008.

1. Wagner, G. P. *et al.* Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* **452**, 470–473 (2008).

- Kenney-Hunt, J. P. *et al.* Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits. *Genetics* **178**, 2275–2288 (2008).
- Turelli, M. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111**, 165–195 (1985).
- Wagner, G. P. The influence of variation and developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.* **1**, 45–66 (1988).
- Orr, H. A. Adaptation and the cost of complexity. *Evolution Int. J. Org. Evolution* **54**, 13–20 (2000).
- Wingreen, N. S., Miller, J. & Cox, E. C. Scaling of mutational effects in models of pleiotropy. *Genetics* **164**, 1221–1228 (2003).
- Martin, G. & Lenormand, T. A general multivariate extension of Fisher's geometrical model and the distribution of mutational effects across species. *Evolution Int. J. Org. Evolution* **60**, 893–907 (2006).
- Orgogozo, V., Broman, K. W. & Stern, D. L. High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two *Drosophila* species. *Genetics* **173**, 197–205 (2006).
- McGregor, A. P. *et al.* Morphological evolution through multiple cis-regulatory mutations at a single gene. *Nature* **448**, 587–590 (2007).
- Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. Two sites in the Delta gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017 (2000).

doi:10.1038/nature07452

Wagner *et al.* reply

Reply to: J. Hermisson & A. P. McGregor *Nature* **456**, doi:10.1038/nature07452 (2008)

In our paper on pleiotropic scaling and the cost of complexity¹, we presented evidence for three findings: first, most genes affect a small number of traits (the degree of pleiotropy is low); second, the total effect of a quantitative trait locus (QTL) increases with the degree of pleiotropy, refuting the constant total effect model^{2,3}; and third, the increase in total effect (defined as $T = \sum_i A_i^2$, where A_i is the effect on

character i , that is, half the difference between the genotypic values of the homozygous genotypes) seems to be stronger than predicted by the superposition model^{4,5} of pleiotropic effects. Hermisson and McGregor⁶ point out that the last result could be due to multiple mutations being mapped to the same QTL, but only if these mutations affect overlapping sets of traits. We agree that this is a possibility that we could not address with the data at hand.

Hermisson and McGregor⁶ also point out that the total effect, if measured using a Manhattan metric, $t = \sum_i |A_i|$, is not affected by double mutations at a QTL, unlike the Euclidian metric we adopted. This is an important observation, as it can be used to estimate the overlap between the set of traits affected by the QTLs and their average pleiotropy. Applying this method to the four QTLs in our data set with the highest degrees of pleiotropy and which deviate

most strongly from the superposition model (see Supplementary Note 6 in ref. 1), we obtain the results shown in Table 1.

It follows from the final column of Table 1 that a double mutation model implies that the mutations at each double mutant-QTL have to have quite high degrees of pleiotropy (about 19 on average), which is unlikely given the distribution of pleiotropic effects in the rest of the data ($P = 0.002$). Hence the double mutation model, although plausible, is not closely compatible with our data. Resolution of the issue of whether we are faced with multiple mutations or a biological effect clearly requires data from more fine-grained QTL mapping experiments (J.M.C. *et al.*, manuscript in preparation). But for now we agree with Hermisson and McGregor⁶ that the superposition model is a reasonable assumption about the pleiotropic scaling of mutation effects.

Günter P. Wagner¹, Jane P. Kenney-Hunt², Mihaela Pavlicev², Joel R. Peck³, David Waxman³ & James M. Cheverud²

¹Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520-8106, USA.

e-mail: gunter.wagner@yale.edu

²Department of Anatomy and Neurobiology, Washington University, St Louis, Missouri 63110, USA.

³Center for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK.

Table 1 | Analysis of the four highly pleiotropic QTLs

\bar{n}	t	\bar{n}	n_{12}	\bar{n}
26	9.19	41	15	21
28	7.35	33	5	17
29	8.35	37	8	19
30	8.69	39	9	20

This analysis assumes that the effects of these QTLs are due to the combined effect of two mutations in each QTL. Here \bar{n} is the number of affected traits, t is the total effect, when measured as Manhattan distance, \bar{n} is the sum of the number of effects of both mutations based on the assumption that the effects per trait are invariant ($\bar{n} = \hat{n}/0.2236$, where the denominator is the regression coefficient of t on \bar{n}), n_{12} is the overlap of the trait sets affected by the two mutations ($n_{12} = \bar{n} - \hat{n}$), and $\bar{n} = \hat{n}/2$ is the average number of traits affected by the two mutations (average degree of pleiotropy).

1. Wagner, G. P. *et al.* Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* **452**, 470–472 (2008).

2. Orr, H. A. Adaptation and the cost of complexity. *Evolution Int. J. Org. Evolution* **54**, 13–20 (2000).

3. Wingreen, N. S., Miller, J. & Cox, E. C. Scaling of mutational effects in models of pleiotropy. *Genetics* **164**, 1221–1228 (2003).

4. Turelli, M. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111**, 165–195 (1985).

5. Wagner, G. P. The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.* **1**, 45–66 (1988).

6. Hermisson, J. & McGregor, A. P. Pleiotropic scaling and QTL data. *Nature* **456**, doi:10.1038/nature07452 (2008).

doi:10.1038/nature07453

Efficient tumour formation by single human melanoma cells

Elsa Quintana^{1*}, Mark Shackleton^{1*}, Michael S. Sabel², Douglas R. Fullen³, Timothy M. Johnson⁴ & Sean J. Morrison¹

A fundamental question in cancer biology is whether cells with tumorigenic potential are common or rare within human cancers. Studies on diverse cancers, including melanoma, have indicated that only rare human cancer cells (0.1–0.0001%) form tumours when transplanted into non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice. However, the extent to which NOD/SCID mice underestimate the frequency of tumorigenic human cancer cells has been uncertain. Here we show that modified xenotransplantation assay conditions, including the use of more highly immunocompromised NOD/SCID interleukin-2 receptor gamma chain null (*Il2rg*^{−/−}) mice, can increase the detection of tumorigenic melanoma cells by several orders of magnitude. In limiting dilution assays, approximately 25% of unselected melanoma cells from 12 different patients, including cells from primary and metastatic melanomas obtained directly from patients, formed tumours under these more permissive conditions. In single-cell transplants, an average of 27% of unselected melanoma cells from four different patients formed tumours. Modifications to xenotransplantation assays can therefore dramatically increase the detectable frequency of tumorigenic cells, demonstrating that they are common in some human cancers.

Traditionally, many cancer cells have been considered to have tumorigenic potential, even though no assay has yet demonstrated that a high percentage of single human cancer cells can form tumours. In contrast, the cancer stem-cell model has suggested that only small subpopulations of cancer cells have tumorigenic potential, based on experiments in which human cancer cells were xenotransplanted into NOD/SCID mice. For example, only one in a million (0.0001%) human melanoma cells is tumorigenic in NOD/SCID mice¹. Indeed, most human cancers have only rare (<0.1%) tumorigenic/leukaemogenic cells (also called cancer-initiating cells or cancer stem cells) when transplanted into NOD/SCID or other highly immunocompromised mice^{1–11}. Nonetheless, recent studies of mouse haematopoietic malignancies have raised the question of whether NOD/SCID assays underestimate the frequency of human cancer-initiating cells^{12–14}. Indeed, human leukaemias exhibit a modestly higher frequency of leukaemogenic cells when assayed in mice that are more highly immunocompromised than NOD/SCID mice^{15,16}, although leukaemogenic cells still represent only 1% of cells in one such model¹⁷. The critical question is whether optimization of xenotransplantation assays could reveal that some human cancers actually have very common cells with tumorigenic potential despite only having rare tumorigenic cells in NOD/SCID mice.

The question of whether cells with tumorigenic potential are common or rare within human cancers has fundamental implications for therapy. If tumorigenic cells represent small minority populations, as suggested by the evidence supporting the cancer stem-cell model, improved anti-cancer therapies may be identified based on the ability to kill these cancer stem cells rather than the bulk population of non-tumorigenic cancer cells^{18,19}. Alternatively, if cells with tumorigenic potential are common, it will not be possible to treat cancer more effectively or to understand cancer biology better by focusing on small minority subpopulations.

Melanoma-initiating cells are rare in NOD/SCID mice

Melanoma-initiating (tumorigenic) cells were reported to be rare based on the observation that only 1 in 1,090,000 human metastatic

melanoma cells formed tumours within 8 weeks of transplantation into NOD/SCID mice¹. To assess this, we transplanted 10²–10⁷ freshly dissociated melanoma cells obtained directly from seven patients subcutaneously into NOD/SCID mice (see Supplementary Table 1 for more information on tumours). Palpable tumours were evident in some mice 8 weeks after injection of cells from four of seven melanomas (Fig. 1A, B). Limiting dilution analysis²⁰ indicated that the average frequency of cells that formed tumours within 8 weeks of transplantation into NOD/SCID mice was 1 in 837,000 (Fig. 1C), confirming the published estimate¹. However, most tumours took more than 8 weeks to develop (Fig. 1A). On average, tumours first became palpable after 11.4 ± 3.8 weeks (mean ± s.d.), or 14.3 ± 7.6 weeks for tumours that arose from fewer than 10,000 injected cells. Variability was high, but the average frequency of cells that formed tumours within 32 weeks was 1 in 111,000 (Fig. 1C; *P* < 0.0001). The frequency of melanoma-initiating cells is therefore significantly underestimated when tumour formation is monitored for only 8 weeks.

Modified assay reveals more tumorigenic cells

Some normal human haematopoietic cells engraft more efficiently in NOD/SCID mice lacking the interleukin-2 gamma receptor (NOD/SCID *Il2rg*^{−/−}) compared with NOD/SCID mice, owing in part to the lack of natural-killer cell activity in NOD/SCID *Il2rg*^{−/−} mice^{21–24}. NOD/SCID *Il2rg*^{−/−} mice have also been used to study cancer arising from human cell lines^{25,26} or human leukaemias^{15,27}. We thus compared human melanoma growth in NOD/SCID mice and NOD/SCID *Il2rg*^{−/−} mice to test whether more tumorigenic cells could be detected in more highly immunocompromised NOD/SCID *Il2rg*^{−/−} mice. Xenografted melanoma cells (human melanomas grown in NOD/SCID mice) from five patients were dissociated, then live human cells were isolated by flow cytometry (excluding mouse haematopoietic and endothelial cells; Fig. 2a) and transplanted side-by-side into NOD/SCID *Il2rg*^{−/−} and NOD/SCID mice (Fig. 2b). Tumours grew faster in NOD/SCID *Il2rg*^{−/−} mice (Fig. 2b and

¹Howard Hughes Medical Institute, Life Sciences Institute, Department of Internal Medicine, and Center for Stem Cell Biology, University of Michigan, Ann Arbor, Michigan 48109-2216, USA. ²Department of Surgery, ³Department of Pathology, and ⁴Department of Dermatology, University of Michigan, Ann Arbor, Michigan 48109, USA.

*These authors contributed equally to this work.

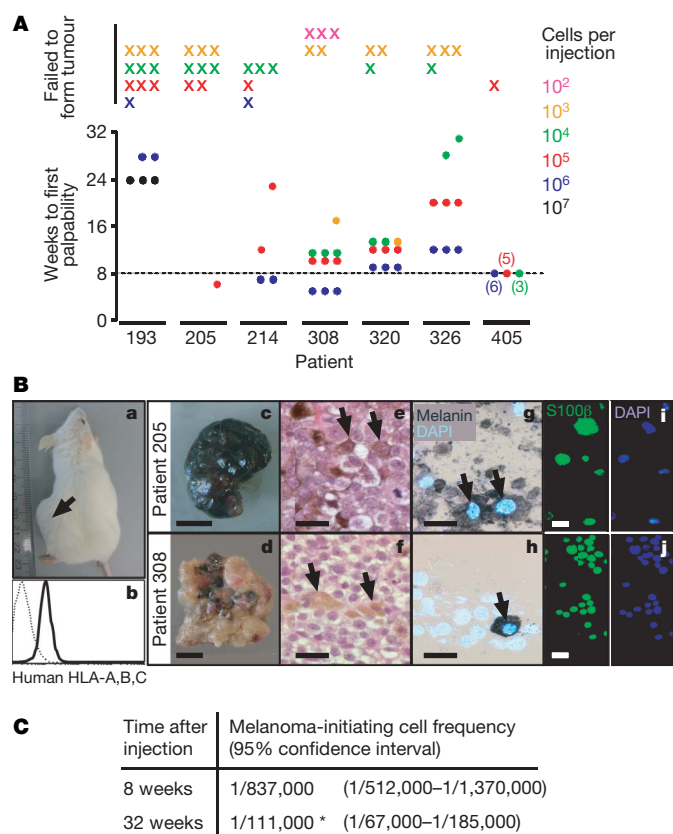


Figure 1 | Only rare human melanoma cells form tumours in NOD/SCID mice. **A**, Tumour development after subcutaneous injection of unfractionated primary melanoma cells directly from seven patients into NOD/SCID mice. Dots represent the times after injection at which individual tumours were first palpable and are coloured according to cell dose. Crosses are injections that failed to form tumours. Dotted line indicates 8 weeks after injection. **B**, All tumours were diagnosed as metastatic melanoma by clinical pathology (see Supplementary Table 1 for more information). The tumours that formed in mice (**a**, arrow) became large, grew quickly once they were palpable and were histologically similar to the patient tumours from which they were derived. Flow cytometry demonstrated that most tumour cells expressed human HLA (**b**; dotted line represents unstained control). Some tumours were highly pigmented (**c**) whereas others contained variable pigmentation (**d**) or were amelanotic (scale bar, 1 cm). Sections stained with haematoxylin and eosin through the same tumours showed pigmented cells (**e**, **f**, see arrows; bars, 25 μ m). Cytospun cells contained melanin, as indicated by Fontana-Masson staining (**g**, **h**, arrows; bars, 25 μ m), and showed widespread S100 β staining (**i**, **j**), a marker used to diagnose melanoma⁴⁰. **C**, Limiting dilution analyses of the frequency of tumorigenic melanoma cells in Fig. 1A at 8 weeks or 32 weeks after transplantation (* $P < 0.0001$).

Supplementary Fig. 1), and an increased ($P < 0.05$) frequency of melanoma-initiating cells was observed in NOD/SCID *Il2rg*^{-/-} mice compared with NOD/SCID mice in every tumour tested (Supplementary Fig. 2). Two melanoma specimens obtained directly from patients (465 and 481) also exhibited a significantly ($P < 0.05$) higher frequency of melanoma-initiating cells in NOD/SCID *Il2rg*^{-/-} mice (Supplementary Fig. 2).

To test whether NOD/SCID *Il2rg*^{-/-} mice selected for growth of more aggressive clones, we transplanted cells from tumours grown in NOD/SCID *Il2rg*^{-/-} mice back into NOD/SCID versus NOD/SCID *Il2rg*^{-/-} recipients. The frequency of detectable tumorigenic cells in these tumours went back down in NOD/SCID mice (Supplementary Fig. 3a), demonstrating that heritable changes in the frequency of melanoma-initiating cells did not occur as a result of having grown in NOD/SCID *Il2rg*^{-/-} mice. An increased xenogeneic immune response probably contributes to reduced tumorigenesis by melanoma cells in NOD/SCID mice compared with NOD/SCID *Il2rg*^{-/-} mice.

We next tested whether further improvements could be made in melanoma cell engraftment. Co-injection with Matrigel²⁸ increases tumour formation by cancer cell lines^{29,30} and enhances the engraftment of primary human epithelial cancer cells in immunocompromised mice^{6,31}. To test the effect of Matrigel on the ability to detect tumorigenic melanoma cells, we isolated live human melanoma cells by flow cytometry (Fig. 2a) from xenografts derived from three patients. The same cell preparations were transplanted into NOD/SCID *Il2rg*^{-/-} mice after mixing with either vehicle (see Methods) or vehicle with Matrigel. Tumour cells injected with Matrigel produced tumours that grew faster than cells injected with vehicle alone (Fig. 2c and Supplementary Fig. 4a), and limiting dilution analysis revealed that in every case more melanoma cells were tumorigenic in Matrigel (Supplementary Fig. 4b).

To quantify the combined effect of the individual assay improvements described above, we performed side-by-side transplantation of live melanoma cells isolated by flow cytometry from xenografted tumours from three different patients. The same cell preparations were either mixed with vehicle and injected into NOD/SCID mice or mixed with Matrigel and injected into NOD/SCID *Il2rg*^{-/-} mice. In every case, we observed much higher frequencies of tumorigenic cells in the NOD/SCID *Il2rg*^{-/-} mice. On average, more than 5,000-fold more cells exhibited tumorigenic activity under the modified assay conditions, in which an average of 1 in 9 human melanoma cells formed tumours (Fig. 2d). These results indicate that xenotransplantation assays can be modified to detect much higher than expected frequencies of human cells with tumorigenic potential.

Tumorigenic potential is common in melanoma

To ensure that our modified assay conditions did not somehow confer tumorigenic capacity on normal human cells, we injected 10,000 primary human melanocytes and/or 100,000 primary human mesenchymal stem cells with Matrigel into NOD/SCID *Il2rg*^{-/-} mice. These injections did not lead to the formation of tumours (Fig. 3a). The frequency of tumorigenic melanoma cells also did not increase with passaging in mice (Supplementary Fig. 4c). Therefore, this xenotransplantation assay does not confer tumorigenic potential on human cells. Consistent with this, serial transplantation experiments demonstrated that the increased tumorigenicity of melanoma cells in NOD/SCID *Il2rg*^{-/-} mice did not reflect a heritable change in the transplanted cells themselves (Supplementary Fig. 3a).

To determine whether xenografted melanomas consistently exhibit high frequencies of tumorigenic cells, we tested five additional tumour samples in an independent series of experiments, including some tumours that exhibited rare tumorigenic cells in NOD/SCID mice (Fig. 1). In each case, live human melanoma cells were isolated by flow cytometry (Fig. 2a) and injected with Matrigel into NOD/SCID *Il2rg*^{-/-} mice. Palpable tumours from the injection of eight cells were first detected 10 ± 2 weeks after injection. Every tumour exhibited a high frequency of tumorigenic cells (range 1 in 21 to 1 in 5 cells), averaging 1 in 9 cells (Fig. 3b).

To assess whether melanoma cells obtained directly from patients also contain a high frequency of cells with tumorigenic potential, we assessed cells obtained from six patients. These included four metastatic melanomas and two primary cutaneous melanomas. Live melanoma cells were isolated by flow cytometry, excluding human haematopoietic (CD45⁺ or glycophorinA⁺) and endothelial cells (CD31⁺), which collectively represented $50 \pm 33\%$ of cells in tumours obtained directly from patients. These excluded cells were confirmed to be haematopoietic and endothelial cells rather than melanoma cells by microscopy, and were greatly depleted of tumorigenic activity when transplanted (data not shown). The melanoma cells isolated by flow cytometry were mixed with Matrigel and injected into NOD/SCID *Il2rg*^{-/-} mice. On average, palpable tumours from the injection of 10 cells were first detected 13 ± 4 weeks after injection. Every tumour exhibited a high frequency of

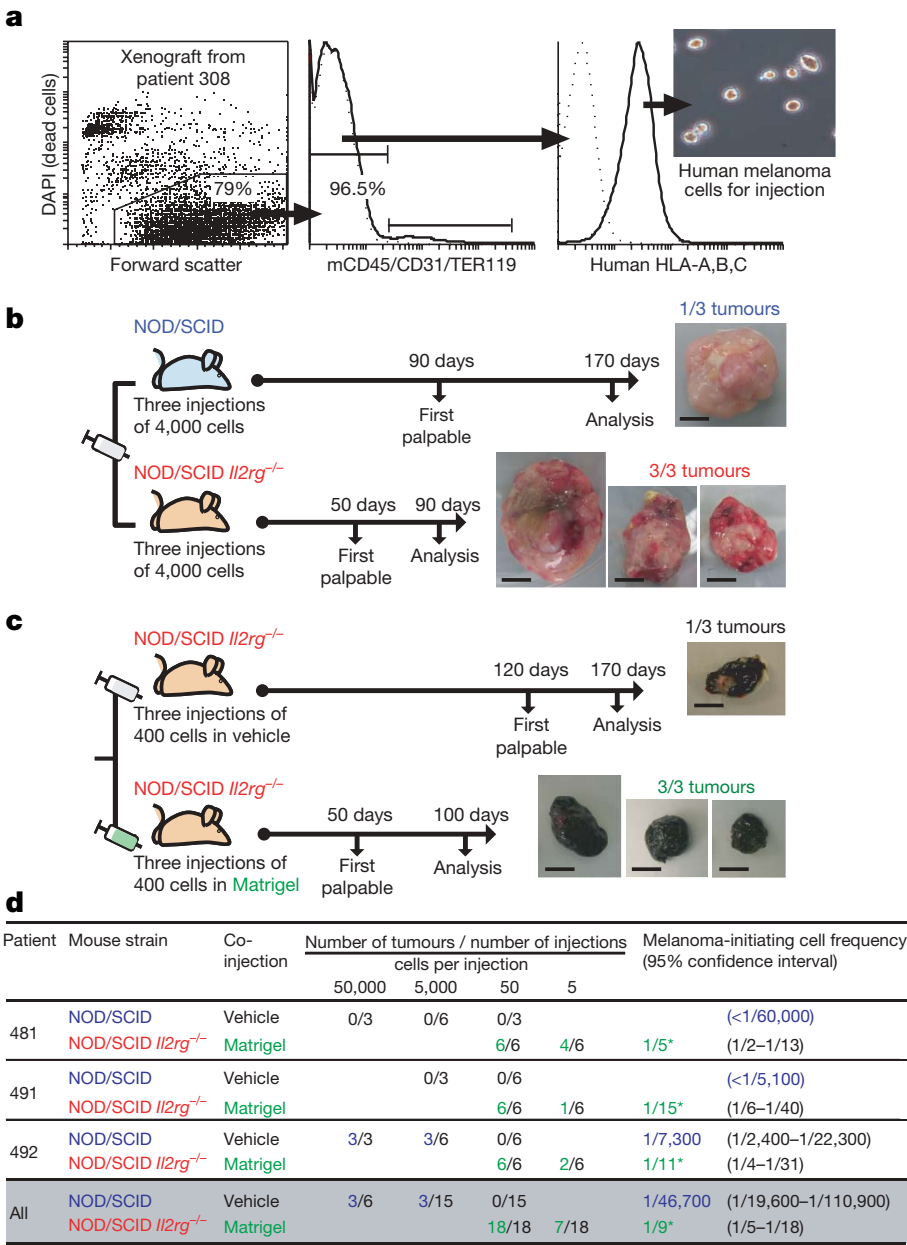


Figure 2 | Modifications to the xenotransplantation assay reveal that many more human melanoma cells have tumorigenic potential than detected in NOD/SCID mice. **a**, Live human melanoma cells were isolated from xenografted tumours by flow cytometry. After gating to eliminate debris and dead cells, additional gates were drawn to select human HLA⁺ cells and to exclude mouse haematopoietic (CD45 and TER119) and endothelial (CD31) cells (middle panel). The human HLA⁺ cells consistently formed tumours upon transplantation into immunocompromised mice whereas mouse haematopoietic and endothelial cells did not (data not shown). **b**, Tumour development after side-by-side subcutaneous injections of 4,000 human melanoma cells from a xenograft derived from patient 308 into NOD/SCID or NOD/SCID Il2rg^{-/-} mice. **c**, Tumour development after side-by-side injections of 400 human melanoma cells from a xenograft derived from patient 205 cells into NOD/SCID Il2rg^{-/-} mice, with or without Matrigel. Photographs in **b** and **c** show resulting tumours at the time of analysis (bars, 1 cm). Similar experiments conducted with xenografted and non-xenografted tumours from several patients demonstrated that a significantly higher frequency of human melanoma cells formed tumours when injected into NOD/SCID Il2rg^{-/-} mice (Supplementary Fig. 2) or with Matrigel (Supplementary Fig. 4b). The tumours were also palpable earlier and grew faster (Supplementary Figs 1 and 4a). **d**, Limiting dilution analyses of tumours that arose after 8 weeks in NOD/SCID or NOD/SCID Il2rg^{-/-} mice from the side-by-side transplantation of melanoma cells (derived from xenografts) mixed with vehicle or Matrigel, respectively. In all cases, melanoma-initiating cells were significantly (**P* < 0.05) more frequent in the modified assay and represented 1 in 5 to 1 in 15 cells.

tumorigenic cells, with an average of 1 in 4 injected cells forming a tumour based on limiting dilution analysis (Fig. 3c). In tumours obtained directly from patients 492, 501 and 509 (the last a cutaneous primary tumour), every injection (6/6) of only 10 cells produced a tumour (Fig. 3c).

We have thus performed injections of small numbers of melanoma cells from 12 different patients and have not yet found a tumour that contained rare tumorigenic cells. Tumorigenic cells were common in all tumours, irrespective of whether they were derived from xenografts or directly from patients, and whether they were from primary cutaneous or metastatic melanomas.

Tumorigenesis by single, unselected melanoma cells

To our knowledge, no study has yet demonstrated that a high percentage of single cells from a spontaneously occurring human cancer has the potential to form tumours *in vivo*. To assess this, we sorted live human melanoma cells by flow cytometry from xenografted tumours obtained from four different patients, then deposited one cell per well in Terasaki plates (well volume 10 μ l). Wells were visually confirmed to contain a single cell, then mixed with Matrigel and injected into NOD/SCID Il2rg^{-/-} mice (Fig. 4a). Depending on

the patient, 12–65% of single cells formed tumours (Fig. 4b). Overall, 69 tumours (27%) developed from 254 single-cell injections. This demonstrates that xenotransplantation assays can be improved to the point where single human melanoma cells can engraft and form tumours *in vivo*, confirming that cells with tumorigenic potential are common within human melanomas.

Additional modifications to xenotransplantation assays may further improve the detection of human cells able to form tumours, such that 25% could remain an underestimate of the percentage of melanoma cells with tumorigenic potential. Because some melanoma cells are fated to undergo cell death or senescence as a result of deleterious genetic changes or proximity to necrotic areas, it is possible that most melanoma cells that are not fated to undergo cell death or senescence have tumorigenic potential. It also remains possible that some melanomas, particularly early-stage primary tumours, may contain less frequent tumorigenic cells than observed in our studies.

Tumorigenic cells are phenotypically heterogeneous

To assess whether tumorigenic melanoma cells are phenotypically distinct from melanoma cells that fail to form tumours, we examined the expression of more than 50 surface markers on melanomas

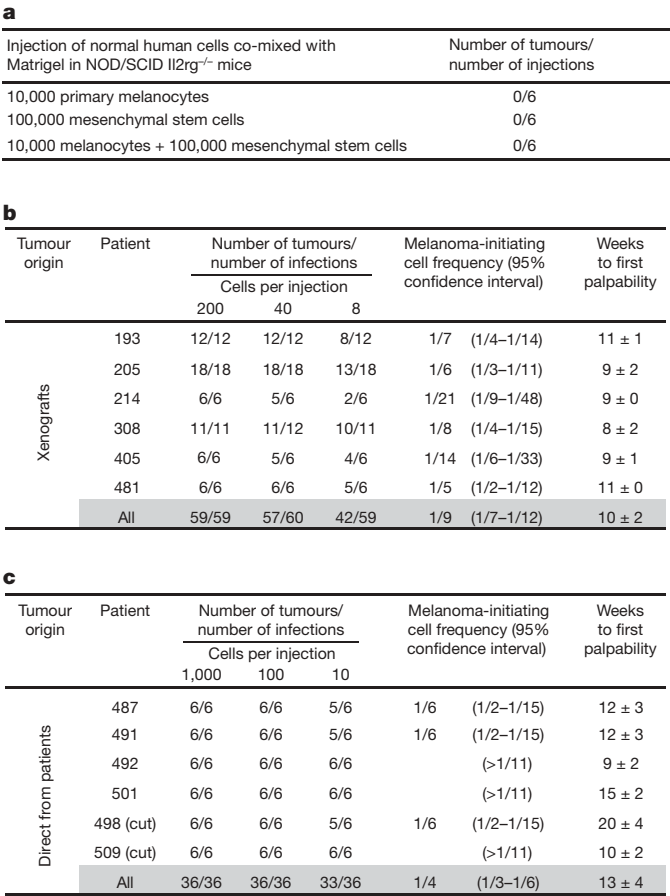


Figure 3 | A high percentage of human melanoma cells are tumorigenic, but normal human cells are not. **a**, Large doses of primary human melanocytes and/or mesenchymal stem cells were mixed with Matrigel and transplanted into NOD/SCID *Il2rg*^{-/-} mice. No tumours were palpable after 22 weeks. The frequency of melanoma-initiating cells in 12 human tumours obtained either from xenografts (**b**) or directly from patients (**c**) was calculated by limiting dilution analysis. Tumours from patients 498 and 509 were primary cutaneous lesions (cut), whereas other tumours were metastases (Supplementary Table 1). In each case, live human melanoma cells were isolated by flow cytometry, mixed in Matrigel and injected subcutaneously into NOD/SCID *Il2rg*^{-/-} mice. Eleven per cent of melanoma cells from xenografts and 25% of melanoma cells directly from patients formed tumours within 32 weeks of transplantation. Weeks to palpability (mean ± s.d.) are indicated for the lowest dose of cells from each tumour.

derived from several patients (Supplementary Table 2). These included markers of other cancer stem-cell populations, melanocytes, melanoma, neural-crest derivatives and other cell types. Fifteen of these markers (A2B5, c-kit, CD44, CD49B, CD49D, CD49f, CD54, CD133, CD166, E-cadherin, HNK-1, L1CAM, MCAM, N-cadherin and p75) were heterogeneously expressed by melanoma cells from multiple patients and were tested for the ability to distinguish tumorigenic from non-tumorigenic melanoma cells *in vivo* (Supplementary Table 2). In each case, melanoma cells were fractionated by flow cytometry (except CD133, which was sometimes fractionated using magnetic beads as in previous studies^{10,32}) and cells that expressed different levels of the indicated markers were injected into NOD/SCID *Il2rg*^{-/-} mice. In every case, tumours arose from all fractions of cells. We found no marker that distinguished tumorigenic from non-tumorigenic cells (Supplementary Table 2). Detailed results are shown for CD49f (α_6 integrin, a marker expressed by many different stem cells³³) and L1CAM (which is associated with CD133 expression in glioma stem cells³⁴) in Supplementary Fig. 5.

A previous study found that 1 in 120,735 (0.00083%) ABCB5⁺ metastatic melanoma cells formed tumours in NOD/SCID mice, a

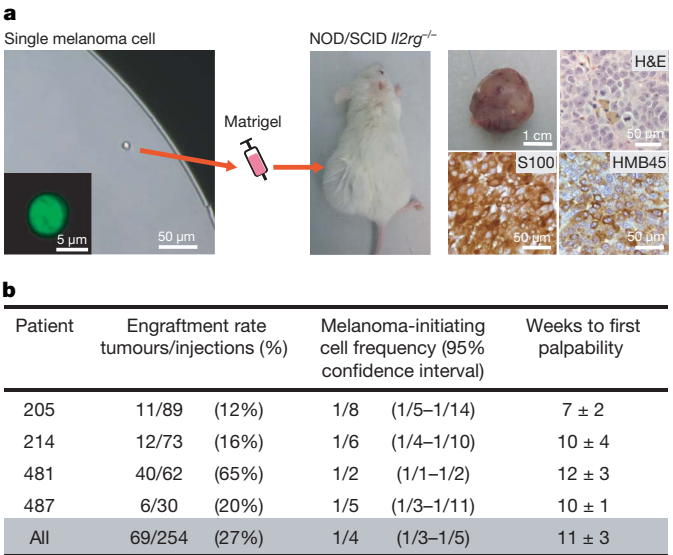


Figure 4 | Efficient tumour development from the xenotransplantation of single human melanoma cells. **a**, Flow-cytometrically isolated human melanoma cells derived from xenografts from four patients were diluted into Terasaki microwells such that wells containing single cells could be identified by phase contrast microscopy. In control experiments, the presence of single cells was confirmed by the observation of single nuclei with Acridine Orange staining (inset) in 90 out of 90 cases. The single cells were mixed with Matrigel and injected into NOD/SCID *Il2rg*^{-/-} mice. Tumours arising from the injection of single cells were confirmed to be melanoma by haematoxylin and eosin, S100 and HMB45 staining (right panels show sections from a tumour that arose from a single cell obtained from patient 214). **b**, The percentage of single-cell injections (69/254 = 27%) that formed tumours within 20 weeks of transplantation. Weeks to first palpability (mean ± s.d.) are indicated for each set of tumours.

tenfold enrichment over unfractionated cells¹. Because ABCB5 expression has been shown to correlate with the expression of CD166 and CD133 (ref. 35), we tested whether CD166 or CD133 could enrich tumorigenic melanoma cells in the modified xenotransplantation assay (Supplementary Fig. 6). The frequency of CD133⁺ cells in tumours from 12 different patients was consistently lower (usually lower than 5%) than the frequency of tumorigenic cells in the same tumours (5–20%; Supplementary Fig. 6b). Moreover, CD133⁺ cells were not enriched for tumorigenic melanoma cells. Both CD133⁺ and CD133⁻ fractions from two different melanomas exhibited very high frequencies of tumorigenic cells (Supplementary Fig. 6c). Flow cytometry indicated that all of the resulting tumours contained both CD133⁺ and CD133⁻ cells, irrespective of whether they were derived from CD133⁺ or CD133⁻ cells (data not shown). Both CD166⁺ and CD166⁻ fractions also contained very high frequencies of tumorigenic melanoma cells (Supplementary Fig. 6f). We have therefore been unable to identify any phenotypic differences that distinguish tumorigenic from non-tumorigenic melanoma cells. These results raise the possibility that markers that enrich rare cells with tumorigenic potential in NOD/SCID mice may fail to distinguish tumorigenic from non-tumorigenic cells in assays that detect much higher frequencies of tumorigenic cells, though more work will be required to test this fully.

Discussion

Our data demonstrate that modifications in xenotransplantation assays can dramatically increase—by several orders of magnitude—the detectable frequency of cells with tumorigenic potential. This means that some cancers that appear to have rare tumorigenic cells in NOD/SCID mice actually have very common cells with tumorigenic capacity under other conditions. Other cancers may still have infrequent tumorigenic cells, even when studied under optimized conditions, but the frequency and phenotypic diversity of these cells

may be considerably greater than currently thought. Efforts to optimize the xenotransplantation of human cancer cells will be necessary to identify and study the full spectrum of human cancer cells capable of contributing to disease progression.

It is important to note that neither our study nor previous cancer stem-cell studies have addressed which cells actually contribute to tumour growth and disease progression in patients. These studies address the potential of cancer cells to proliferate extensively and to form tumours/leukaemias, not their actual fate within patients. Depending on how different the tumour environment is within patients compared with mouse models, it is possible that different cancer cells form tumours in mouse models than in human patients. Thus, although we observe a high percentage of melanoma cells that have the potential to proliferate extensively and form new tumours, it is possible that an even greater, or a much smaller, fraction of melanoma cells actually contributes to disease progression in patients.

Although most cancer stem-cell studies have detected only rare cells with tumorigenic capacity, it has recently been pointed out that leukaemogenic/tumorigenic cells do not necessarily have to be rare for cancers to follow a cancer stem-cell model¹⁷. Our observation that NOD/SCID xenotransplantation can dramatically underestimate the frequency of tumorigenic cells in at least some human cancers does not necessarily mean that such cancers will not have intrinsically different populations of tumorigenic and non-tumorigenic cells. Having said that, the frequency of tumorigenic cells in human melanoma is much higher than reported for any cancer previously suggested to follow a cancer stem-cell model, and we have not yet been able to identify phenotypic differences between melanoma cells that form tumours and those that do not. If markers that distinguish tumorigenic from non-tumorigenic melanoma cells in optimized xenotransplantation assays are identified in future, it will be important to test whether they reflect epigenetic differences between cells (as envisioned under the cancer stem-cell model) or genetic/environmental differences between cells.

Although cells with tumorigenic potential are likely to be much more frequent in most human cancers than estimated based on NOD/SCID transplantation, we believe the available evidence continues to support the conclusion that some human cancers follow a cancer stem-cell model. For example, data from the syngeneic transplantation of mouse acute myeloid leukaemia (AML)³⁶ and the transplantation of human AMLs into improved mouse models¹⁷ continue to suggest that many AMLs have small, intrinsically distinct subpopulations of AML-initiating cells. Extensive clinical experience with germ-cell cancers proves that therapies that eliminate the undifferentiated subset of cancer cells can cure patients, even if differentiated cancer cells are left behind^{37,38}. Nonetheless, careful optimization of xenotransplantation assays will probably yield a more nuanced view in which some cancers contain small subpopulations of cancer-initiating cells, whereas others contain common tumorigenic cells with little evidence of hierarchical organization. In both cases it will be critical to identify all cancer cells that have the potential to contribute to disease in patients in order to develop more effective therapies.

METHODS SUMMARY

Tumour cell preparation. Melanoma specimens were obtained from patients according to protocols approved by the Institutional Review Board of the University of Michigan Medical School (IRB MED approval numbers 2004-1058 and 2000-0713). Fresh tumour tissue was mechanically dissociated, enzymatically digested and filtered to obtain a single-cell suspension.

Cell labelling and sorting. Cells were stained with directly conjugated antibodies to human CD45, human CD31 and glycophorin A (to eliminate haematopoietic and endothelial cells from tumours obtained directly from patients) or mouse CD45, mouse CD31, Ter119 and human leukocyte antigen (HLA)-A, -B, -C (to eliminate mouse haematopoietic and endothelial cells and select human cells from xenografted tumours). Cells were resuspended in $10 \mu\text{g ml}^{-1}$ 4,6-diamidino-2-phenylindole (DAPI) to label dead cells and sorted on FACS Vantage SE or FACS Aria flow cytometers. Some samples were labelled with anti-CD133 antibody and separated magnetically using a CD133 Cell Isolation

Kit (Miltenyi Biotec). When testing other markers, cells were usually labelled with unconjugated primary antibodies (see Supplementary Table 2) and then with directly conjugated secondary antibodies before staining with the directly conjugated antibodies described above.

Transplantation of human melanoma cells. After sorting, cells were counted and re-suspended in staining medium (L15 medium containing 1 mg ml^{-1} BSA, 1% penicillin/streptomycin and 10 mM HEPES (pH7.4)), with or without 25% Matrigel (BD Biosciences). Subcutaneous injections of human melanoma cells were performed in NOD.CB17-Prkdcscid/J (NOD/SCID) and NOD.CB17-Prkdcscid Il2rgtm1Wjl/SzJ (NOD/SCID *Il2rg*^{-/-}) mice (Jackson Laboratories) according to protocols approved by the Committee on the Use and Care of Animals at the University of Michigan (protocol number 9055).

Statistics. Limiting dilution analyses were performed based on Bonnefoix *et al.*³⁹, using the limdil function of the 'StatMod' package (author G.K. Smyth, <http://bioinf.wehi.edu.au/software/limdil/>), part of the R statistical software project (<http://www.r-project.org>). Melanoma-initiating cell frequencies were compared using likelihood ratio tests.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 August; accepted 14 October 2008.

- Schatton, T. *et al.* Identification of cells initiating human melanomas. *Nature* **451**, 345–349 (2008).
- Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030–1037 (2007).
- Prince, M. E. *et al.* Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl Acad. Sci. USA* **104**, 973–978 (2007).
- Wu, C. *et al.* Side population cells isolated from mesenchymal neoplasms have tumor initiating potential. *Cancer Res.* **67**, 8216–8222 (2007).
- Wang, J. C. *et al.* High level engraftment of NOD/SCID mice by primitive normal and leukemic hematopoietic cells from patients with chronic myeloid leukemia in chronic phase. *Blood* **91**, 2406–2414 (1998).
- Al-Hajj, M. *et al.* Prospective identification of tumorigenic breast cancer cells. *Proc. Natl Acad. Sci. USA* **100**, 3983–3988 (2003).
- Cox, C. V. *et al.* Characterization of acute lymphoblastic leukemia progenitor cells. *Blood* **104**, 2919–2925 (2004).
- Hope, K. J., Jin, L. & Dick, J. E. Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. *Nature Immunol.* **5**, 738–743 (2004).
- Dalerba, P. *et al.* Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl Acad. Sci. USA* **104**, 10158–10163 (2007).
- O'Brien, C. A. *et al.* A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106–110 (2007).
- Ricci-Vitiani, L. *et al.* Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111–115 (2007).
- Williams, R. T., den Besten, W. & Sherr, C. J. Cytokine-dependent imatinib resistance in mouse BCR-ABL⁺, *Arf*-null lymphoblastic leukemia. *Genes Dev.* **21**, 2283–2287 (2007).
- Kelly, P. N. *et al.* Tumor growth need not be driven by rare cancer stem cells. *Science* **317**, 337 (2007).
- Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).
- Agliano, A. *et al.* Human acute leukemia cells injected in NOD/LtSz-SCID/IL- γ ^{null} mice generate a faster and more efficient disease compared to other NOD/SCID-related strains. *Int. J. Cancer* **123**, 2222–2227 (2008).
- Feuring-Buske, M. *et al.* Improved engraftment of human acute myeloid leukemia progenitor cells in β -2-microglobulin-deficient NOD/SCID mice and in NOD/SCID mice transgenic for human growth factors. *Leukemia* **17**, 760–763 (2003).
- Kennedy, J. A. *et al.* Comment on 'Tumor growth need not be driven by rare cancer stem cells'. *Science* **318**, 1722; author reply 1722 (2007).
- Pardal, R., Clarke, M. F. & Morrison, S. J. Applying the principles of stem-cell biology to cancer. *Nature Rev. Cancer* **3**, 895–902 (2003).
- Wang, J. C. & Dick, J. E. Cancer stem cells: lessons from leukemia. *Trends Cell Biol.* **15**, 494–501 (2005).
- Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
- Shultz, L. D. *et al.* Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R γ ^{null} mice engrafted with mobilized human hemopoietic stem cells. *J. Immunol.* **174**, 6477–6489 (2005).
- Ito, M. *et al.* NOD/SCID/ γ _c^{null} mouse: an excellent recipient mouse model for engraftment of human cells. *Blood* **100**, 3175–3182 (2002).
- McKenzie, J. L. *et al.* Human short-term repopulating stem cells are efficiently detected following intrafemoral transplantation into NOD/SCID recipients depleted of CD122⁺ cells. *Blood* **106**, 1259–1261 (2005).
- Shultz, L. D. *et al.* Regulation of human short-term repopulating cell (STRC) engraftment in NOD/SCID mice by host CD122⁺ cells. *Exp. Hematol.* **31**, 551–558 (2003).

25. Hamada, K. *et al.* Liver metastasis models of colon cancer for evaluation of drug efficacy using NOD/Shi-scid IL2R γ^{null} (NOG) mice. *Int. J. Oncol.* **32**, 153–159 (2008).
26. Suemizu, H. *et al.* Identification of a key molecular regulator of liver metastasis in human pancreatic carcinoma using a novel quantitative model of metastasis in NOD/SCID/ γ^{null} (NOG) mice. *Int. J. Oncol.* **31**, 741–751 (2007).
27. Ishikawa, F. *et al.* Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nature Biotechnol.* **25**, 1315–1321 (2007).
28. Kleinman, H. K. & Martin, G. R. Matrigel: basement membrane matrix with biological activity. *Semin. Cancer Biol.* **15**, 378–386 (2005).
29. Pretlow, T. G. *et al.* Transplantation of human prostatic carcinoma into nude mice in Matrigel. *Cancer Res.* **51**, 3814–3817 (1991).
30. Sweeney, T. M. *et al.* Basement membrane and the SIKVAV laminin-derived peptide promote tumor growth and metastases. *Cancer Metastasis Rev.* **10**, 245–254 (1991).
31. Fridman, R. *et al.* Enhanced tumor growth of both primary and established human and murine tumor cells in athymic mice after coinjection with Matrigel. *J. Natl Cancer Inst.* **83**, 769–774 (1991).
32. Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396–401 (2004).
33. Iwashita, T. *et al.* Hirschsprung disease is linked to defects in neural crest stem cell function. *Science* **301**, 972–976 (2003).
34. Bao, S. *et al.* Targeting cancer stem cells through L1CAM suppresses glioma growth. *Cancer Res.* **68**, 6043–6048 (2008).
35. Frank, N. Y. *et al.* ABCB5-mediated doxorubicin transport and chemoresistance in human malignant melanoma. *Cancer Res.* **65**, 4320–4333 (2005).
36. Yilmaz, O. H. *et al.* Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* **441**, 475–482 (2006).
37. Horwich, A., Shipley, J. & Huddart, R. Testicular germ-cell cancer. *Lancet* **367**, 754–765 (2006).
38. Kleinsmith, L. J. & Pierce, G. B. Multipotentiality of single embryonal carcinoma cells. *Cancer Res.* **24**, 1544–1551 (1964).
39. Bonnefoix, T. *et al.* Fitting limiting dilution experiments with generalized linear models results in a test of the single-hit Poisson assumption. *J. Immunol. Methods* **194**, 113–119 (1996).
40. Nakajima, T. *et al.* Immunohistochemical demonstration of S100 protein in malignant melanoma and pigmented nevus, and its diagnostic application. *Cancer* **50**, 912–918 (1982).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Howard Hughes Medical Institute and by the Allen H. Blondy Research Fellowship. The University of Michigan Melanoma Bank was supported by a gift from Lewis and Lillian Becker. Flow cytometry was partly supported by the University of Michigan Comprehensive Cancer Center grant from the National Institutes of Health CA46592. We thank: D. Adams, M. White and the University of Michigan Flow Cytometry Core Facility for support; N. McAnsh and the University of Michigan Cancer Centre Histology Core for histological studies; G. K. Smyth for assistance with statistics; and Z. Azizan for support with tissue collection. Antibody production was supported in part by the National Institute of Diabetes, Digestive, and Kidney Diseases, grant NIH5P60-DK20572 to the Michigan Diabetes Research and Training Center. Some antibodies were provided by Caltag or by eBioscience to screen for cancer stem-cell markers. Human primary melanocyte cultures were provided by M. Soengas. Human mesenchymal stem cells were provided by Z. Wang and P. Krebsbach. E.Q. was supported by the Spanish Ministry of Education and the Marie Curie Outgoing International Fellowship from the European Commission. M.S. was supported by the Australian National Health and Medical Research Council, the Human Frontiers Science Program and Australia Post.

Author Contributions E.Q., M.S. and S.J.M. planned the project. E.Q. and M.S. performed the experiments, and analysed data with S.J.M. M.S.S. and T.M.J. obtained consent from the patients and surgically obtained many of the melanoma specimens. D.R.F. performed all pathology and diagnosed the tumours with T.M.J. T.M.J. banked the melanomas, and provided clinical information. E.Q., M.S. and S.J.M. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.J.M. (seanjm@umich.edu).

METHODS

Cell preparation. Tumours were mechanically dissociated with a McIlwain tissue chopper (Mickle Laboratory Engineering Co.) before sequential enzymatic digestion in 200 U ml⁻¹ collagenase IV (Worthington) for 20 min followed by 0.05% trypsin-EGTA for 5 min, both at 37 °C. DNase (50–100 U ml⁻¹) was added to reduce clumping of cells during digestion. Cells were filtered (40-µm cell strainer) to obtain a single-cell suspension. Dead cells and debris were reduced by density centrifugation (1.1 g ml⁻¹ Optiprep, Sigma) and/or magnetic bead separation (Dead Cell Removal Kit; Miltenyi Biotec) as necessary. Primary human melanocytes⁴¹ and mesenchymal stem cells⁴² were cultured as described.

Cell labelling and sorting. All antibody labelling of cells was performed for 20 min on ice, followed by washing and centrifugation. Secondary antibodies were conjugated to phycoerythrin (goat anti-mouse IgG or IgM, goat anti-rat IgG or goat anti-rabbit IgG; Jackson ImmunoResearch). Primary isotype controls followed by the same secondary antibodies were used to set background. Cells were subsequently stained with directly conjugated antibodies to human CD45 (HI30-APC, BD Biosciences), human CD31 (WM59-APC, eBiosciences) and Glycophorin A (HIR2-APC, Biolegend) (for tumours obtained directly from patients) or mouse CD45 (30-F11-APC, eBiosciences), mouse CD31 (390-APC, Biolegend), Ter119 (TER-119-APC, eBiosciences) and human HLA-A,B,C (G46-2.6-FITC, BD Biosciences) (for xenograft tumours) to select live human melanoma cells and to exclude endothelial and haematopoietic cells. Cells were re-suspended in 10 µg ml⁻¹ DAPI (Sigma) and sorted on a FACSVantage SE-dual laser, three line flow cytometer or a FACSaria Cell Sorter (Becton Dickinson). After sorting, an aliquot of sorted cells was always re-analysed to check for purity, which was usually greater than 95%. Magnetic cell separation using the CD133 Cell Isolation Kit (Miltenyi Biotec) was performed according to the manufacturer's instructions.

Identification of single melanoma cells. Sorted cells were diluted and aliquoted into 10-µl microwells (Thermo Fisher Scientific). Plates were centrifuged at 450g for 30 s and wells containing single cells were identified by phase-contrast microscopy. In control experiments, the presence of microscopically identified single cells was confirmed by Acridine Orange staining, which demonstrated a single nucleus in 90/90 cases. Cell doublets could be identified easily, were rare (7 of 312 wells examined) and were discarded. Wells containing single cells were transferred to a syringe containing Matrigel before injection. Cell transfer to the syringe was confirmed by observing the absence of a cell in the well it came from.

Transplantation of melanoma cells. Subcutaneous injections were performed into each flank and the interscapular region of each mouse. Tumour formation was evaluated regularly by palpation of injection sites, and tumour diameters were measured with callipers. Mice were monitored for up to 32 weeks after injection. In cases where a tumour became palpable at only one injection site and was causing the mouse distress, that tumour was surgically removed to allow continued evaluation of other injection sites. Tumours were only evident at injection sites and we never observed subcutaneous metastases. To confirm this,

we performed a single subcutaneous injection of 200 cells from tumour 205 into each of six NOD/SCID *Il2rg*^{-/-} mice after mixing in Matrigel and observed only one tumour per mouse at the injection site. The presence of human melanomas was always confirmed at necropsy based on gross examination and marker expression.

Histopathology and immunostaining. Portions of melanoma tumours used in experiments were fixed in 10% neutral buffered formalin, paraffin embedded, sectioned and stained with haematoxylin and eosin for histopathology analysis. Paraffin-embedded tumours were confirmed as melanomas by staining for S100 and HMB45 expression after quenching endogenous peroxidase activity. Binding of anti-S100 antibody (DAKO) was performed for 30 min at room temperature, detected by anti-rabbit secondary (30 min at room temperature) and revealed using DAB chromagen. Binding of HMB45 antibody (DAKO) was performed for 30 min at room temperature after antigen retrieval with proteinase K, detected using the M.O.M. immunodetection kit (Vector Laboratories) and revealed using DAB. S100- and HMB45-stained slides were counterstained with haematoxylin. For staining of sorted cells, cells from tumours were cytopun (18g for 6 min) onto slides after fixation with 4% paraformaldehyde for 5 min at room temperature. For immunofluorescence, slides were rinsed in PBS and blocked in goat serum solution (PBS containing 5% goat serum, 1% BSA and 0.3% Triton X-100) for 1 h to reduce non-specific antibody binding. Incubation with S-100β antibody (Sigma, diluted 1:200 in goat serum solution) was performed overnight at 4 °C, followed by secondary goat anti-mouse IgG1 Alexa 488 (Invitrogen) for 2 h at room temperature. Slides were counterstained with DAPI for 10 min at room temperature, then mounted in fluorescent mounting solution (DAKO). For detection of melanin pigment, the Fontana-Masson method was used⁴³. Slides were microwaved for 1 min in 2.5% Fontana silver nitrate solution (Sigma), before rinsing and toning with 0.2% gold chloride for 2 min. After rinsing, slides were incubated in 5% sodium thiosulphate (Sigma) for 2 min, washed and nuclei were stained with DAPI. Slides were mounted in DAKO fluorescent mounting solution.

Statistics. Differences between mean times to tumour palpability were compared using unpaired *t*-tests. Tumour growth rates were determined by calculating averaged linear regression slopes of the diameters of each tumour at each time point, monitored for at least 10 days and displayed with dot points representing the mean (±s.d.) of the diameters of all tumours palpable at the indicated time points.

41. Fernandez, Y. *et al.* Differential regulation of noxa in normal melanocytes and melanoma cells by proteasome inhibition: therapeutic implications. *Cancer Res.* 65, 6294–6304 (2005).
42. Wang, Z. *et al.* Ablation of proliferating marrow with 5-fluorouracil allows partial purification of mesenchymal stem cells. *Stem Cells* 24, 1573–1582 (2006).
43. Bancroft, J. D. & Stevens, A. *Theory and Practice of Histological Techniques* (Churchill Livingstone, 1990).

Centrosome misorientation reduces stem cell division during ageing

Jun Cheng^{1*}, Nezaket Türkel^{2*†}, Nahid Hemati^{2*}, Margaret T. Fuller⁴, Alan J. Hunt¹ & Yukiko M. Yamashita^{2,3}

Asymmetric division of adult stem cells generates one self-renewing stem cell and one differentiating cell, thereby maintaining tissue homeostasis. A decline in stem cell function has been proposed to contribute to tissue ageing, although the underlying mechanism is poorly understood. Here we show that changes in the stem cell orientation with respect to the niche during ageing contribute to the decline in spermatogenesis in the male germ line of *Drosophila*. Throughout the cell cycle, centrosomes in germline stem cells (GSCs) are oriented within their niche and this ensures asymmetric division. We found that GSCs containing misoriented centrosomes accumulate with age and that these GSCs are arrested or delayed in the cell cycle. The cell cycle arrest is transient, and GSCs appear to re-enter the cell cycle on correction of centrosome orientation. On the basis of these findings, we propose that cell cycle arrest associated with centrosome misorientation functions as a mechanism to ensure asymmetric stem cell division, and that the inability of stem cells to maintain correct orientation during ageing contributes to the decline in spermatogenesis. We also show that some of the misoriented GSCs probably originate from dedifferentiation of spermatogonia.

Adult stem cell populations maintain highly differentiated but short-lived cells such as blood, intestinal epithelium cells and sperm throughout life. Upon division of stem cells, daughter cells must either self-renew to preserve stem cell identity or commit to differentiation. The balance between stem cell self-renewal and differentiation is critical to tissue homeostasis, with disruption of this balance leading to tumorigenesis (caused by stem cell overproliferation) or tissue degeneration (caused by stem cell depletion). To maintain this critical balance, many stem cells have the potential to divide asymmetrically, producing one daughter stem cell and one differentiating cell¹. Many stem cells reside in a special microenvironment, or stem cell niche, that regulates stem cell maintenance^{2,3}. Asymmetric stem cell division within the niche essentially relies on the correct placement of daughter cells inside and outside the niche: daughter cells that remain within the niche retain a stem cell identity whereas daughter cells displaced from the niche are fated to differentiate⁴. Thus, it is critical to establish stem cell polarity within the context of the niche.

A decline in the function of adult stem cells has been proposed to contribute to tissue ageing, although the underlying mechanisms remain enigmatic⁵. Tissue ageing has been proposed to have arisen as a tumour suppressor mechanism^{6,7}, in which tumour suppressor activity reduces stem cell function in later stages of life, preventing tumorigenesis but reducing tissue regenerative capacity⁸. However, the cellular and molecular basis of such phenomena is poorly understood. Although cell cycle inhibitors such as Ink4a are known to accumulate in stem cells with age and to contribute to an age-related decline in tissue regenerative capacity^{9–11}, the mechanisms regulating increased expression of cell cycle inhibitors and their relationship to normal stem cell function are unknown.

Male GSCs in *Drosophila melanogaster* always undergo asymmetric division. The divisions are regulated by a combination of signal(s) from the niche and spindle orientation. The hub cells, which constitute

the stem cell niche, secrete the signalling ligand Unpaired (Upd), which activates the JAK–STAT (Janus kinase–signal transducer and activator of transcription) pathway in the neighbouring germ cells to promote stem cell maintenance^{12,13}. The spindle lies perpendicular to the hub, so that one daughter cell inherits the attachment to the hub, while the other is displaced away from it¹⁴. This stereotypical orientation of the mitotic spindle is precisely set by the positioning of the centrosomes during interphase. The mother centrosome is always anchored to the hub–GSC interface, whereas the daughter centrosome migrates towards the opposite side of the GSC¹⁵. In this way, GSCs are oriented with respect to the niche throughout the cell cycle. Recently, similar centrosome behaviour and stem cell polarity have been reported in the *Drosophila* neuroblast, suggesting that centrosome orientation within stem cells has a general role in asymmetric division^{16,17}. However, the relative importance of this orientation to physiological stem cell function is unclear.

Spermatogenesis declines with age in *Drosophila* testis

Testes from newly eclosed males contain cells in all stages of spermatogenesis. These include transit-amplifying cells (gonialblast and 2–16-cell spermatogonia), spermatocytes, meiotic cells and elongated spermatids (Fig. 1a, b), the collective presence of which indicates ongoing spermatogenesis. In contrast, as flies age, testes undergo marked involution and the number of early germ cells in the apical region of the testis (spermatogonia, spermatocytes and meiotic cells) progressively decreases (Fig. 1b, arrow). A decrease in spermatogenesis could be attributable to decreased function of GSCs. However, the number of GSCs (defined as germ cells attached to the hub) did not significantly decrease after 30 days of age (Fig. 1c, d), when testes had already undergone significant involution. This suggests that decreased stem cell number does not fully explain the reduced spermatogenesis that is observed at this stage of ageing, although stem cell numbers ultimately decline in older flies (50 days)

¹Department of Biomedical Engineering, Center for Ultrafast Optical Science. ²Life Sciences Institute, Center for Stem Cell Biology, ³Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109, USA. ⁴Departments of Developmental Biology and Genetics, Stanford University, School of Medicine, Stanford, California 94305, USA. [†]Present address: Peter MacCallum Cancer Centre, Anatomy and Cell Biology Department, University of Melbourne, Melbourne, Victoria 3002, Australia.

*These authors contributed equally to this work.

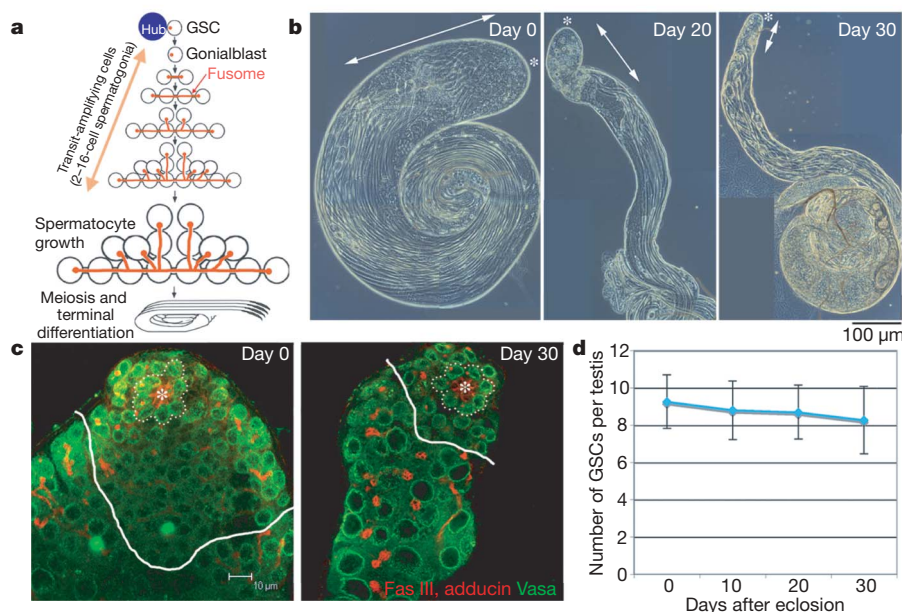


Figure 1 | *Drosophila* testis undergoes an age-related decline in spermatogenesis. **a**, Spermatogenesis of *Drosophila melanogaster* (adapted from ref. 27). GSCs are supported by the hub cells. Each spermatogonial division is incomplete, and the resultant spermatogonia and spermatocytes are connected by a cytoplasmic bridge, or a ring canal, through which a branched fusome runs. **b**, Phase microscopy of ageing testes. The apical

(asterisk) area containing round, relatively early germ cells (arrow) decreases over time. **c**, The number of GSCs (surrounded by dotted line) remains constant with age. White lines separate spermatogonia and spermatocytes. Red, fasciclin III (Fas III; hub) and adducin (fusome); green, Vasa (germ cells). The hub is indicated by an asterisk. **d**, The number of GSCs is shown (\pm s.d.). $n > 50$ testes were counted for each time point.

and may account for decreased spermatogenesis in later stages of ageing¹⁸ (Supplementary Table 1).

The ageing testis accumulates misoriented GSCs

In *Drosophila* male GSCs the stereotypical orientation of the mitotic spindle is set up by the positioning of centrosomes during interphase (Fig. 2a)^{14,15}. In young males (0–2 days after eclosion), GSCs remain oriented towards the niche throughout the cell cycle as previously reported, setting up the stereotypically oriented mitotic spindle. However, we found that, in aged flies, GSCs in which neither centrosome was situated next to the hub became more numerous (hereafter referred to as ‘misoriented centrosomes’, Fig. 2b). GSCs with misoriented centrosomes increased to approximately 40% of total GSCs at 30 days after eclosion (Fig. 2c, d). It should be noted that owing to the scoring criteria (Fig. 2b), the maximum misorientation would be $\sim 50\%$ (as each randomly positioned centrosome would end up adjacent to the hub $\sim 25\%$ of time). Consistent with this, the frequency of misoriented GSCs reached a plateau at $\sim 40\%$, with a similar misorientation frequency in very old flies (day 42–50, Supplementary Table 1).

Misoriented GSCs divide less frequently

Although $\sim 40\%$ of GSCs had misoriented centrosomes in aged testes, we rarely observed ‘misoriented spindles’ in which neither spindle pole is associated with the hub (Fig. 2b), suggesting that GSCs with misoriented centrosomes do not enter mitosis. Overall, 95–100% of mitotic spindles were oriented perpendicular to the hub throughout mitosis, regardless of age (Fig. 3a, b). Not only did misoriented GSCs fail to undergo mitosis, but GFP–PACT (pericentrin/AKAP450 C-terminal domain) labelling of centrosomes revealed that significantly fewer cells with misoriented GSCs underwent the G1/S transition relative to those with oriented GSCs (Fig. 3c; see Methods). Notably, the percentage of oriented GSCs from aged flies (day 10 to 30) that entered the S phase remained as high as that of oriented GSCs from young flies, demonstrating that aged GSCs do not divide less frequently if their centrosomes are appropriately oriented. In contrast, misoriented GSCs entered the S phase with a significantly lower

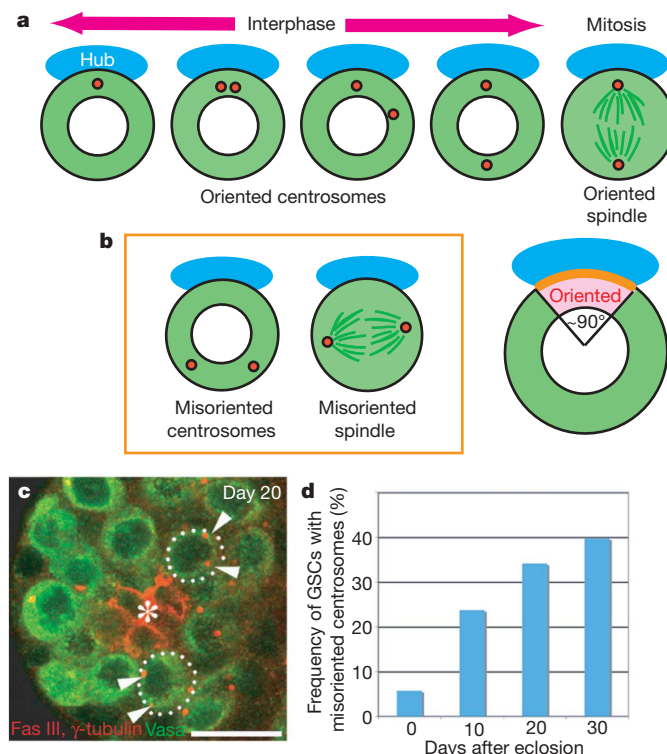


Figure 2 | Misoriented GSCs increase with age. **a**, Schematic diagram of centrosome positioning during the cell cycle. **b**, Left panel: the definition of misoriented centrosomes and spindle. Right panel: scoring criterion. Centrosomes were scored to be oriented when one of two centrosomes is in the pink area close to the hub–GSC junction (orange line). **c**, An example of testis (20-day-old) containing GSCs with misoriented centrosomes. Red, fasciclin III and γ -tubulin (centrosome); green, Vasa. The hub is indicated by an asterisk. Scale bar, 10 μ m. **d**, The frequency of GSCs with misoriented centrosomes increases with age ($n > 275$ GSCs for each time point). The same trend was observed in more than three separate experiments, including conditions with different culture media and temperature (22 $^\circ$ C to 25 $^\circ$ C).

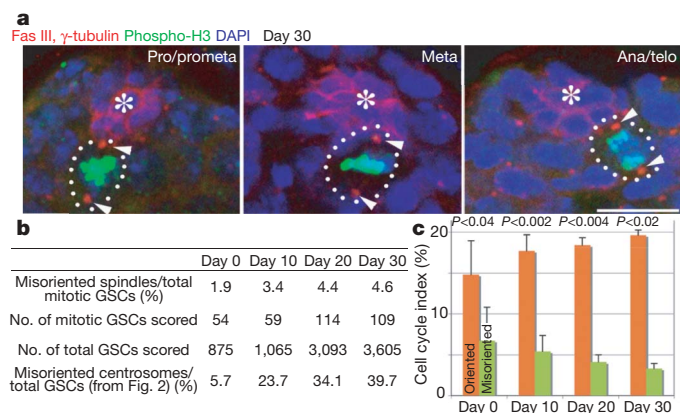


Figure 3 | Misoriented GSCs divide less frequently compared with oriented GSCs. **a**, Spindles remain oriented throughout mitosis even at day 30. Red, fasciclin III and γ -tubulin; green, Thr 3-phosphorylated histone H3 (phospho-H3; mitotic chromatin); blue, DAPI. The hub is indicated by an asterisk. Scale bar, 10 μ m. Arrowheads indicate mitotic spindle poles. **b**, The frequency of misoriented spindles remained low for up to 30 days of age. The percentage of misoriented interphase cells was taken from Fig. 2c for comparison. **c**, Pulse-labelling of centrosomes by heat-shock-induced GFP-PACT expression confirmed the low cell cycle activity of misoriented GSCs. The percentage of GSCs committed to the G1/S transition during the labelling period is shown (mean \pm s.d.; $n > 110$ labelled GSCs for each point, equivalent to $n > 750$ total GSCs per data point). P values of t -test (two-tail) are shown.

frequency compared with oriented GSCs at all stages of life. These results suggest that the reduced division of ageing GSCs is determined by centrosome misorientation.

Misoriented GSCs maintain a capability of cell division

The rare division of misoriented GSCs raises the question of whether misorientation precedes cell cycle arrest/delay or whether it occurs as a consequence of general disintegration of stem cell activity, such as from cellular senescence. To test whether misoriented GSCs retain the ability to re-enter the cell cycle, GSCs were labelled by feeding flies with 5-bromodeoxyuridine (BrdU)-containing food from day 5 to 10, during which time misorientation increased from approximately 10% to 20%. This labelled 96% of all GSCs, including almost all of the misoriented GSCs (which accounted for \sim 20% of all GSCs; Supplementary Fig. 1a). This demonstrates that GSCs that are misoriented by day 10 had replicated their DNA in the preceding several days. Furthermore, when BrdU was discontinued, BrdU labelling quickly disappeared during the BrdU chase period, and was completely absent after 5 days (Supplementary Fig. 1b), indicating that misoriented BrdU-positive GSCs had divided. The ability of mis-oriented GSCs to dilute their BrdU label indicates that misoriented GSCs are not permanently arrested in the cell cycle.

Time-lapse imaging of centrosomes in GSCs confirmed that mis-oriented GSCs did not divide, but were able to resume dividing as soon as centrosomes restored their correct orientation (Fig. 4). We tracked centrosome behaviour within GSCs by using mCherry-Sas6 (ref. 16), and confirmed that most GSCs exhibit stereotypical movement of centrosomes during the cell cycle as shown in Fig. 2a. In addition, we recorded 18 misoriented GSCs that underwent mitosis during the recording time. In each case, misoriented centrosomes moved around within GSCs for a long time (on average, 137 ± 47 min from the start point of the recording), but as long as no centrosome aligned adjacent to the hub the cells never divided. However, when one centrosome became very close to the hub-GSC interface (or 're-oriented'), the GSCs quickly divided (in 18 ± 12 min after re-orientation). The ability of GSCs to re-enter the cell cycle almost immediately on reorientation of the centrosomes suggests that centrosome misorientation is the proximal cause of the reduced GSC division and thus the testis involution observed during fly ageing.

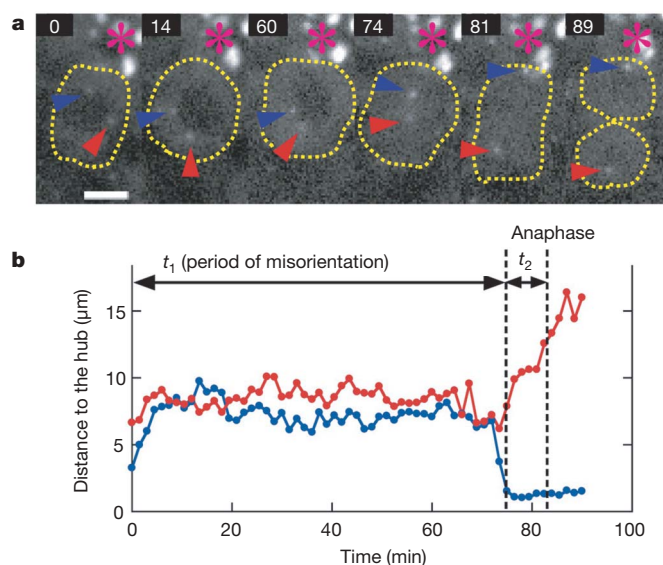


Figure 4 | Misoriented GSCs correct centrosome orientation before mitosis. **a**, Frames from time-lapse imaging of GSC cell division with misoriented centrosome (frames were taken from Supplementary Movie 1). Centrosome movement in a GSC was tracked by mCherry-Sas6. Numbers represent minutes from the start point of recording. Re-orientation occurred around 74 min, and anaphase around 81 min. The hub is indicated by an asterisk; two centrosomes are indicated by arrowheads (red and blue). Scale bar, 5 μ m. **b**, Tracking of the centrosome movement from the same time-lapse imaging. Red and blue lines represent the movement of centrosomes with red and blue arrowheads shown in **a**, respectively.

Some misoriented GSCs originate from dedifferentiation

Dedifferentiation has been proposed to be one mechanism by which the stem cell pool can be replenished^{19,20}, although it is not clear how physiologically important this mechanism is in unperturbed tissues or how it is affected by ageing. We hypothesized that if spermatogonia dedifferentiate to stem cells, such cells might not be able to orient towards the hub initially. To test the hypothesis that misoriented GSCs may originate from dedifferentiated spermatogonia, we used FLP-recombinase-based mitotic recombination to mark permanently germ cells once committed to differentiation (Fig. 5a). FLP recombinase was expressed under the control of *bam*-Gal4, which is selectively expressed in differentiating cells at the 4-cell spermatogonia stage and later²¹. FLP activity in these cells recombined *Actin* promoter>FRT-stop-FRT>LacZ-NLS to generate a β -galactosidase expression in the differentiating cells (*Actin*-LacZ-NLS). In testes from third instar larvae, weak LacZ expression was apparent in 4–8-cell spermatogonia and early spermatocytes (Fig. 5b, between dashed and solid lines), whereas LacZ expression became much stronger in later spermatocytes (Fig. 5b, after solid line). This pattern of LacZ expression persisted for 50 days. Notably, we also sometimes observed LacZ expression in GSCs, as seen by LacZ staining next to the hub (Fig. 5c, arrowhead). The frequency of LacZ-positive (LacZ^+) stem cells increased with age, with 0.7% GSCs being positive in third instar larvae, 6.1% in newly eclosed males and 40% in 50-day-old males (Supplementary Fig. 2). These LacZ^+ GSCs appear to result from dedifferentiation of spermatogonia.

To test whether the LacZ expression in GSCs might be attributable to random, spurious activation of the *bam* promoter in GSCs, we tested whether X-irradiation affected LacZ expression in GSCs. Because irradiation kills GSCs (as well as other cell types), this should increase the rate of dedifferentiation, whereas if LacZ expression is attributable to random *bam* activation in GSCs it should not be affected by irradiation. The frequency of LacZ^+ GSCs increased after X-irradiation, suggesting that LacZ expression in GSCs reflects the

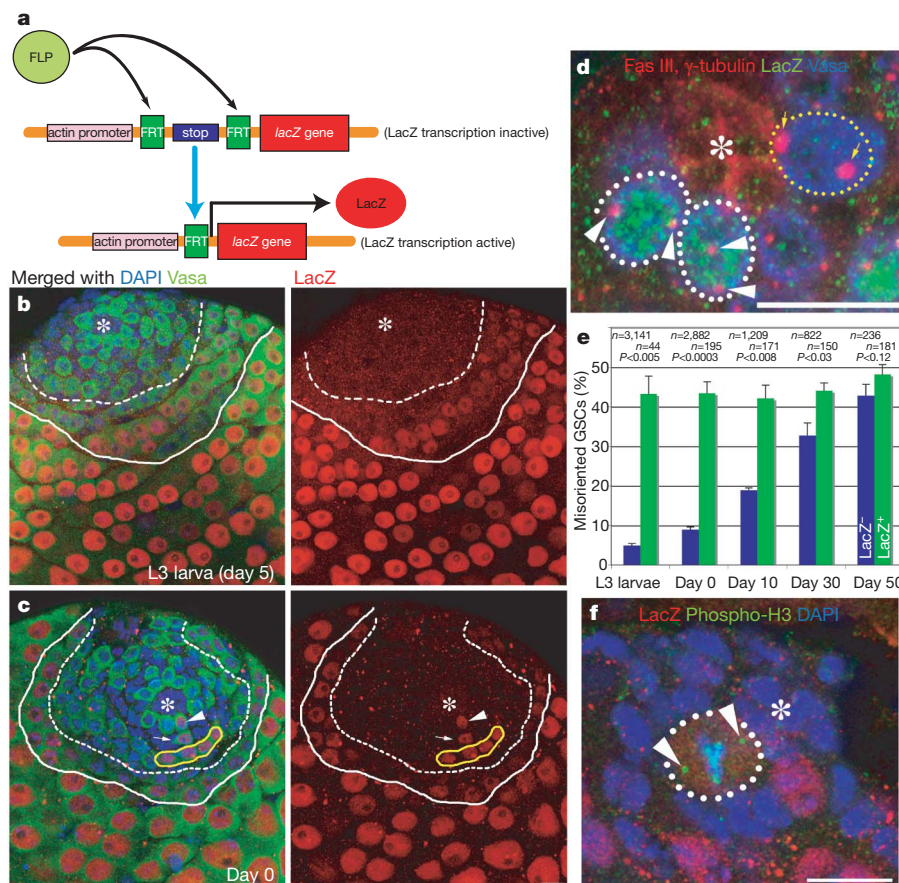


Figure 5 | Dedifferentiated GSCs have a higher frequency of centrosome misorientation. **a**, Strategy to label germ cells that committed to differentiation program through FLP-based recombination. **b**, **c**, LacZ staining of testes from an L3 larva (**b**) and a day 0 adult (**c**). Weak LacZ expression starts around 4–8-cell spermatogonia (dashed line) and becomes stronger in spermatocytes (solid line). The arrowhead in **c** shows a LacZ⁺ GSC with its progeny (gonialblast is shown by an arrow and 4 cell spermatogonia is surrounded by a yellow line). Red, LacZ; green, Vasa; blue, DAPI. The hub is indicated by an asterisk. **d**, An example of testis with two

misoriented LacZ⁺ GSCs (white dotted line, with arrowheads indicating misoriented centrosomes) and a correctly oriented LacZ⁻ GSC (yellow dotted line, with arrows indicating properly oriented centrosomes). Red, fasciclin III and γ -tubulin; green, LacZ; blue, Vasa. **e**, LacZ⁺ GSCs have a high frequency of misorientation (mean \pm s.d.). *P* values of *t*-test (two-tail) are shown. **f**, LacZ⁺ GSCs are correctly oriented during mitosis. Fourth chromosomes that segregate before anaphase mark spindle poles (arrowheads). Red, LacZ; green, phospho-H3; blue, DAPI. Scale bars in **c**, **d**, **f** indicate 10 μ m.

replenishment of GSCs by dedifferentiation of spermatogonia (Fig. 6a). In further support of this conclusion, the frequency that multiple LacZ⁺ GSCs were next to each other (Fig. 6b) was significantly higher than what would be expected based on the random, sporadic expression of FLP in GSCs (Fig. 6a), indicating simultaneous dedifferentiation of interconnected spermatogonia. Furthermore, we observed LacZ⁺ spermatogonia next to the hub with disintegrating fusomes and ring canals (Fig. 6c and Supplementary Fig. 3), hallmarks of dedifferentiation^{19,20} (0.11 dedifferentiating events per testis after 2,000 or 3,000 rad of irradiation, *n* = 388). Notably, we also observed many LacZ-negative (LacZ⁻) dedifferentiating spermatogonia (Fig. 6d and Supplementary Fig. 3, 0.47 dedifferentiating events per testis, *n* = 388), suggesting that not all dedifferentiation events are marked by LacZ expression (on average, only ~18% of total dedifferentiation events were marked by LacZ expression). This is presumably due to the dedifferentiation from 2–4-cell spermatogonia that have not activated the *bam* promoter. Taken together, these results suggest that LacZ expression successfully marks dedifferentiation.

Notably, we found that LacZ⁺ GSCs had a significantly higher frequency of centrosome misorientation compared to LacZ⁻ GSCs (Fig. 5d, e). From L3 larvae to day 50 after eclosion, >40% of LacZ⁺ GSCs showed centrosome misorientation, suggesting that dedifferentiated GSCs cannot correctly orient centrosomes towards the hub. It should be noted that not all misoriented GSCs were LacZ⁺. The

frequency of misoriented LacZ⁻ GSCs increased with age, reaching ~30% after 30 days and ~40% after 50 days. These GSCs might result from dedifferentiation of gonialblasts and 2–4-cell spermatogonia, which have not yet expressed *bam*-induced FLP recombinase, as suggested above. Alternatively, there might be other reasons that cause centrosome misorientation in GSCs that remain to be determined. In either case, these results demonstrate that misoriented GSCs originate, at least in part, from the dedifferentiation of spermatogonia.

Importantly, the mitotic spindles of dedifferentiated GSCs were oriented similarly to constitutive (LacZ⁻) GSCs (Fig. 5f; 97% oriented (*n* = 202) for LacZ⁻ GSCs and 95% oriented (*n* = 20) for LacZ⁺ GSCs). This demonstrates that, whereas centrosome orientation is defective in dedifferentiated GSCs, those that resume division exhibit normal mitotic spindle orientation. These data indicate that the dedifferentiated GSCs that enter mitosis do so only when their centrosomes are oriented. Considering that we rarely observed misoriented spindles, even in very early prophase (Fig. 3a), it is unlikely that the misoriented spindle (not the centrosome) undergoes re-orientation after entering mitosis, as observed in *Drosophila* embryonic neuroblast²². This also demonstrates that dedifferentiated GSCs can function as stem cells by dividing and producing progeny, which was also demonstrated by the presence of LacZ⁺ differentiating progenies produced by LacZ⁺ GSCs (Fig. 5c; arrow, gonialblast; yellow solid line, 4-cell spermatogonia).

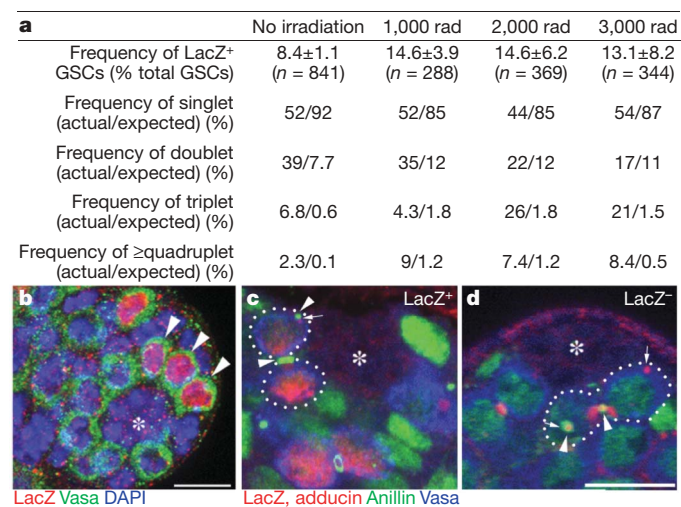


Figure 6 | Evidence that LacZ⁺ GSCs are generated by dedifferentiation **a**, Summary of LacZ⁺ GSC frequencies with or without X-ray irradiation. Flies were irradiated at day 0 of age, and analysed at day 5. The actual and expected frequencies of singlet, doublet, triplet and quadruplet (or more) are shown. *n* indicates the number of GSCs scored. The expected frequencies were calculated as described in Methods. **b**, An example of testis apical tip (irradiated with 2,000 rad) containing a triplet (arrowheads: three LacZ⁺ GSCs are locating next each other). Red, LacZ; green, Vasa; blue, DAPI. Scale bar, 10 μm. The hub is indicated by an asterisk. **c**, **d**, Examples of LacZ⁺ (**c**) and LacZ⁻ (**d**) dedifferentiating spermatogonia (surrounded by dotted lines), with multiple ring canals (arrowheads) and disintegrating fusomes (arrows), observed 24 h after irradiation. Red, LacZ and adducin; green, anillin (ring canal and nucleus); blue, Vasa. Scale bar, 10 μm.

Discussion

Here we demonstrate that GSCs with misoriented centrosomes accumulate as flies age. Because such misoriented GSCs divide less frequently as compared with oriented GSCs, accumulation of misoriented GSCs contributes to the decline in spermatogenesis that occurs with age. Although misoriented GSCs rarely divide, they are not permanently arrested (or senescent) and are correctly oriented when they divide. Whether correction of GSC orientation is an active process that is part of the acquisition of stem cell identity remains to be determined. The low cell cycle activity of misoriented GSCs may also suggest that mechanisms are in place to detect misorientation and induce cell cycle arrest in response to this change, although the underlying mechanisms remain to be identified.

We have also demonstrated that misoriented GSCs originate, at least in part, from dedifferentiation of spermatogonia. Although dedifferentiated GSCs have a high frequency (>40%) of centrosome misorientation they can function as stem cells by resuming the cell cycle, with correctly oriented mitotic spindles just like constitutive GSCs. GSC numbers do not decrease as quickly as expected from the calculated GSC half-life, suggesting that a mechanism to compensate for the loss of GSCs exists²³. Because we rarely observed misoriented spindles, or symmetric stem cell division, we speculate that dedifferentiation is the major mechanism to replace stem cells over time in the *Drosophila* male germ line.

A decline in GSC number in older males (day 50) was reported recently¹⁸. This decrease in stem cell number is probably due to failure of the niche function (via decreased signal from the niche as well as decreased E-cadherin-based attachment between the niche and GSCs)¹⁸. However, the decrease in the production of spermatogonia and testis involution precede the loss of GSCs such that decreasing GSC numbers cannot explain the testis involution that is observed at younger ages.

The present results provide a novel mechanistic link between the control of stem cell polarity and the age-related decline in tissue

regenerative capacity. Mechanisms responsible for monitoring stem cell orientation with respect to the niche not only prevent overproliferation of stem cells by ensuring the asymmetric outcome of the stem cell division, but they contribute to the decline in tissue regenerative capacity during ageing. Many of the misoriented GSCs originate from the dedifferentiation of spermatogonia, a mechanism thought to be responsible for maintaining the stem cell population over extended periods of time. Therefore, although GSCs produce fewer progeny over time, the system appears to maximize the number of progeny produced throughout life, while maintaining asymmetric stem cell division.

We propose that the GSCs with misoriented centrosomes divide less frequently and that a combination of such a decreased stem cell division and a higher frequency of the GSC misorientation in aged testes leads to a decline in spermatogenesis with age.

METHODS SUMMARY

Immunofluorescent staining. Immunofluorescent staining was performed as described elsewhere¹⁴. For γ-tubulin and LacZ dual staining, testes were fixed in 90% ethanol 3.8% formaldehyde solution (chilled to -20 °C). Fixed testes were then permeabilized by washing in 1× PBS with 0.1% Triton X-100 (30 min) before immunofluorescent staining. The following primary antibodies were used in combination with appropriate Alexa-Fluor-conjugated secondary antibodies (1:200, Molecular Probes): mouse anti-fasciclin III (1:10, developed in ref. 24 and obtained from the Developmental Studies Hybridoma Bank), mouse anti-adducin (1:20, developed in ref. 25 and obtained from the Developmental Studies Hybridoma Bank), rabbit anti-Vasa (1:2,000, a gift from R. Lehmann), goat anti-Vasa (1:20, Santa Cruz), mouse anti-γ-tubulin (1:100, GTU-88, Sigma), rabbit anti-phosphorylated histone H3 (Thr3) (1:200, Upstate), mouse anti-β-galactosidase (1:200, G4644, Sigma), rabbit anti-β-galactosidase (1:500, Abcam), rabbit anti-anillin (1:1,300, a gift from C. Field) and mouse anti-BrdU (1:200, BU-33, Sigma). BrdU staining was performed as detailed in ref. 26. Images were taken using a Leica SP5 confocal microscope and processed using Adobe Photoshop CS 8.0.

Time-lapse live imaging of GSCs inside testes. Newly eclosed Sas6-mCherry flies¹⁶ were dissected inside *Drosophila* culture medium containing Schneider's *Drosophila* medium and 10% fetal bovine serum. The testis tips were placed inside a sterile glass-bottom chamber, and was mounted on a three-axis computer-controlled piezoelectric stage and imaged using an inverted microscope equipped with an electron multiplier cooled CCD camera. Four-dimensional image sequences (*x*, *y*, *z* and time) were acquired every 90 s. A semi-automatic tracking software was used to track locations of centrosomes and hubs, and the tracking algorithm was based on the pattern matching routine inside National Instrument IMAQ Vision software. The Supplementary Movie was generated using Sony Vegas software.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 April; accepted 1 September 2008.

Published online 15 October 2008.

- Morrison, S. J. & Kimble, J. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* **441**, 1068–1074 (2006).
- Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414**, 98–104 (2001).
- Fuchs, E., Tumber, T. & Guasch, G. Socializing with the neighbors: stem cells and their niche. *Cell* **116**, 769–778 (2004).
- Yamashita, Y. M., Fuller, M. T. & Jones, D. L. Signaling in stem cell niches: lessons from the *Drosophila* germline. *J. Cell Sci.* **118**, 665–672 (2005).
- Rando, T. A. Stem cells, ageing and the quest for immortality. *Nature* **441**, 1080–1086 (2006).
- Sharpless, N. E. & Depinho, R. A. How stem cells age and why this makes us grow old. *Nature Rev. Mol. Cell Biol.* **8**, 703–713 (2007).
- Serrano, M. & Blasco, M. A. Cancer and ageing: convergent and divergent mechanisms. *Nature Rev. Mol. Cell Biol.* **8**, 715–722 (2007).
- Kirkwood, T. B. Understanding the odd science of aging. *Cell* **120**, 437–447 (2005).
- Krishnamurthy, J. et al. p16^{INK4a} induces an age-dependent decline in islet regenerative potential. *Nature* **443**, 453–457 (2006).
- Molofsky, A. V. et al. Increasing p16^{INK4a} expression decreases forebrain progenitors and neurogenesis during ageing. *Nature* **443**, 448–452 (2006).
- Janzen, V. et al. Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16^{INK4a}. *Nature* **443**, 421–426 (2006).
- Tulina, N. & Matunis, E. Control of stem cell self-renewal in *Drosophila* spermatogenesis by JAK-STAT signaling. *Science* **294**, 2546–2549 (2001).

13. Kiger, A. A., Jones, D. L., Schulz, C., Rogers, M. B. & Fuller, M. T. Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. *Science* **294**, 2542–2545 (2001).
 14. Yamashita, Y. M., Jones, D. L. & Fuller, M. T. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. *Science* **301**, 1547–1550 (2003).
 15. Yamashita, Y. M., Mahowald, A. P., Perlin, J. R. & Fuller, M. T. Asymmetric inheritance of mother versus daughter centrosome in stem cell division. *Science* **315**, 518–521 (2007).
 16. Rusan, N. M. & Peifer, M. A role for a novel centrosome cycle in asymmetric cell division. *J. Cell Biol.* **177**, 13–20 (2007).
 17. Rebollo, E. *et al.* Functionally unequal centrosomes drive spindle orientation in asymmetrically dividing *Drosophila* neural stem cells. *Dev. Cell* **12**, 467–474 (2007).
 18. Boyle, M., Wong, C., Rocha, M. & Jones, D. L. Decline in self-renewal factors contributes to aging of the stem cell niche in the *Drosophila* testis. *Cell Stem Cell* **1**, 470–478 (2007).
 19. Brawley, C. & Matunis, E. Regeneration of male germline stem cells by spermatogonial dedifferentiation *in vivo*. *Science* **304**, 1331–1334 (2004).
 20. Kai, T. & Spradling, A. Differentiating germ cells can revert into functional stem cells in *Drosophila melanogaster* ovaries. *Nature* **428**, 564–569 (2004).
 21. Schulz, C. *et al.* A misexpression screen reveals effects of bag-of-marbles and TGF β class signaling on the *Drosophila* male germ-line stem cell lineage. *Genetics* **167**, 707–723 (2004).
 22. Kaltschmidt, J. A., Davidson, C. M., Brown, N. H. & Brand, A. H. Rotation and asymmetry of the mitotic spindle direct asymmetric cell division in the developing central nervous system. *Nature Cell Biol.* **2**, 7–12 (2000).
 23. Wallenfang, M. R., Nayak, R. & DiNardo, S. Dynamics of the male germline stem cell population during aging of *Drosophila melanogaster*. *Aging Cell* **5**, 297–304 (2006).
 24. Patel, N. H., Snow, P. M. & Goodman, C. S. Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in *Drosophila*. *Cell* **48**, 975–988 (1987).
 25. Ding, D., Parkhurst, S. M. & Lipshitz, H. D. Different genetic requirements for anterior RNA localization revealed by the distribution of Adducin-like transcripts during *Drosophila* oogenesis. *Proc. Natl Acad. Sci. USA* **90**, 2512–2516 (1993).
 26. Gonczy, P. & DiNardo, S. The germ line regulates somatic cyst cell proliferation and fate during *Drosophila* spermatogenesis. *Development* **122**, 2437–2447 (1996).
 27. Fuller, M. T. Spermatogenesis. In *The Development of Drosophila melanogaster* (eds Bate, M. & Martinez-Arias, A.) 71–147 (Cold Spring Harbor Laboratory Press, 1993).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** We thank C. Gonzalez, D. McKearin, N. Rusan, M. Peifer and the Bloomington Stock Center for fly stocks; R. Lehmann, C. Field and the Developmental Studies Hybridoma Bank for antibodies; M. Kiel and D. Nakada for help with X-ray irradiation; and S. Morrison and T. Mahowald for comments on the manuscript. This research was supported by a University of Michigan start-up fund, March of Dimes Basil O'Connor Starter Scholar Research Award and the Searle Scholar Program (to Y.M.Y.), and NIH grants P01 DK53074 (to M.T.F.) and R01GM072006 (to A.J.H.).
- Author Contributions** Y.M.Y. designed research. J.C. and A.J.H. designed and conducted time-lapse imaging of centrosome behaviour. J.C., N.T., N.H. and Y.M.Y. performed other experiments. M.T.F. contributed to research design for Figs 1 and 2. Y.M.Y. wrote the manuscript.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.M.Y. (yukikomy@umich.edu).

METHODS

Fly strains. Heat-shock (hs)-Gal4 (Flybase), Actin<FRT-stop-FRT<LacZ and UAS-FLP flies were obtained from the Bloomington *Drosophila* Stock Center. UAS-GFP-PACT has been described previously¹⁵. *bam*-Gal4 is a gift from D. McKearin. mCherry-Sas6 is a gift from N. Rusan and M. Peifer¹⁶. All *Drosophila* stocks were maintained under standard culture conditions using Bloomington Standard media or yeast-glucose media.

Cell cycle index assessed by GFP-PACT incorporation into centrosome. Heat-shock (hs)-Gal4/UAS-GFP-PACT flies (0, 10, 20, 30 days of age) were subjected to 2.5 h of heat shock at 37 °C. Cells that passed through the G1/S transition during the period of GFP-PACT expression incorporated GFP-PACT into their centrosomes. At 10 h after the heat shock, when most of the GSCs are still in the first cell cycle, flies were dissected and testes were subjected to immunofluorescent staining. Cell cycle index was calculated as GFP-positive GSCs/total GSCs (%).

Pulse-chase of BrdU labelling. *yw* or *Asl-YFP* flies (a gift from C. Gonzalez) were fed from 6 days of age to 10 days of age with BrdU in the mixture of 1 ml 100% apple juice, 0.8 g of instant fly food (Sigma) and 40 µl of 100 mg ml⁻¹ BrdU solution (in 1:1 of acetone and DMSO). Flies were transferred to fresh BrdU-containing food every 24 h. After the feeding period, flies were transferred to regular fly food and subjected to immunofluorescent staining at appropriate time points (day 10 (0-day chase), day 12 (2-day chase) and day 15 (5-day chase)). Some flies did not incorporate BrdU into gonad; thus such testes without any BrdU incorporation were not scored.

Detection of germ cells that have committed to the differentiation program. Male UAS-FLP; *bam*-Gal4 flies were crossed to female Actin<FRT-stop-FRT<LacZ-NLS flies. Progenies were aged appropriately and subjected to immunofluorescent staining for LacZ and other appropriate markers. For X-ray irradiation, flies were irradiated with a caesium 137 GammaCell40 Exactor Irradiator (MDS Nordia), delivering approximately 100 rad min⁻¹.

It should be noted that we occasionally observed LacZ⁺ signal in testicular somatic cells (that is, cyst progenitor cells, cyst cells and hub cells), although much less frequently than LacZ⁺ GSCs. In contrast to LacZ⁺ GSCs, of which frequency increases with age, the frequency of the LacZ⁺ somatic cells did not increase over time (L3 larvae to 30 days of age), suggesting that somatic cells from late larval stage to 30 days of age do not newly create active *lacZ* gene. Although we do not know the reason why *lacZ* gene activation occurs in somatic cells, such events appear to occur only during early stages of development but not after the L3 larval stage.

Statistical analysis of LacZ⁺ GSC frequency. With or without X-irradiation, the frequency of LacZ⁺ GSCs (x) comes from actual observation and is calculated as the number of counted LacZ⁺ GSCs divided by the total number of observed GSCs. Then, the frequency of LacZ⁻ GSCs is $1 - x$. A general statistical analysis is discussed below, excluding unrealistic cases of $x = 0$ or 1 (that is, all GSCs are LacZ⁻ or LacZ⁺).

The singlet is defined as a single LacZ⁺ GSC with adjacent LacZ⁻ GSCs on both sides, and the doublet is defined as two LacZ⁺ GSCs next to each other with adjacent LacZ⁻ GSCs on both sides, and so on. In Fig. 6, the expected frequency was calculated as the 'conditional probability' with the assumption of random LacZ expression in GSCs. Because $x \neq 0$ or 1 , the probability of locating an alignment of LacZ⁺ next to LacZ⁻ GSC (that is, [LacZ⁻, LacZ⁺]) is unity. Under this condition, if the next GSC is LacZ⁻ (with the probability $1 - x$), it is counted as a singlet (that is, [LacZ⁻, LacZ⁺, LacZ⁻]). Thus, the probability of singlet is $1 - x$. If an alignment of [LacZ⁻, LacZ⁺] is followed by a LacZ⁺ GSC (with the probability x) and then a LacZ⁻ GSC (with the probability $1 - x$), it is counted as a doublet (that is, [LacZ⁻, LacZ⁺, LacZ⁺, LacZ⁻]), resulting in the probability of doublet as $(1 - x)x$. Similarly, the probability of triplet (that is, [LacZ⁻, LacZ⁺, LacZ⁺, LacZ⁺, LacZ⁻]) is $(1 - x)x^2$, and the probability of quadruplet or more is $(1 - x)x^3 + (1 - x)x^4 + \dots = x^3$. Here, we did not consider the fact that GSCs are aligned in a circle around the hub, because the probability of all GSCs surrounding the hub (~9 GSCs per testis) being LacZ⁺ is negligible for either actual observations or calculated expected values ($<10^{-7}$). Obviously, the normalization of the total probability is also conserved by summing all probability together as $(1 - x) + (1 - x)x + (1 - x)x^2 + x^3 = 1$.

Mitofusin 2 tethers endoplasmic reticulum to mitochondria

Olga Martins de Brito¹ & Luca Scorrano^{1,2}

Juxtaposition between endoplasmic reticulum (ER) and mitochondria is a common structural feature, providing the physical basis for intercommunication during Ca^{2+} signalling; yet, the molecular mechanisms controlling this interaction are unknown. Here we show that mitofusin 2, a mitochondrial dynamin-related protein mutated in the inherited motor neuropathy Charcot–Marie–Tooth type IIa, is enriched at the ER–mitochondria interface. Ablation or silencing of *mitofusin 2* in mouse embryonic fibroblasts and HeLa cells disrupts ER morphology and loosens ER–mitochondria interactions, thereby reducing the efficiency of mitochondrial Ca^{2+} uptake in response to stimuli that generate inositol-1,4,5-trisphosphate. An *in vitro* assay as well as genetic and biochemical evidences support a model in which mitofusin 2 on the ER bridges the two organelles by engaging in homotypic and heterotypic complexes with mitofusin 1 or 2 on the surface of mitochondria. Thus, mitofusin 2 tethers ER to mitochondria, a juxtaposition required for efficient mitochondrial Ca^{2+} uptake.

Mitochondria regulate crucial cellular events, from Ca^{2+} signalling to apoptosis^{1,2}. To achieve these complex tasks, they are spatially and functionally organized in a network of dynamically interconnected organelles often in close contact with the ER³. This dictates mitochondrial uptake of Ca^{2+} released from the ER by inositol-1,4,5-trisphosphate (InsP_3)^{3–5}, ultimately affecting organelle metabolism⁶ and the response to Ca^{2+} -dependent death stimuli⁷. Furthermore, contacts are crucial for mitochondrial lipid biosynthesis, which occurs at ‘mitochondria-associated membranes’ (MAMs)—patches of ER membranes attached to mitochondria, containing several phospholipid and glycosphingolipid biosynthetic enzymes⁸.

The molecular details of ER–mitochondria tethers are largely unknown, but they are probably proteinaceous⁹. PACS2, a multifunctional sorting protein mainly localized at the ER, indirectly controls juxtaposition of the organelles through BAP31 (also known as BCAP31)-dependent fission and perinuclear clustering of mitochondria¹⁰. Similarly, the dynamin-related GTPase DRP1 (also known as DNM1L), required for fission of mitochondria and ER, can alter tethering by causing perinuclear clumping of the former^{11,12}. Other mitochondrial dynamin-related GTPases include optic atrophy 1 (OPA1), involved in mitochondrial fusion and apoptosis^{13,14}, and mitofusin (MFN) 1 and 2 (ref. 15). MFN1 is crucial for mitochondrial docking and fusion¹⁶, whereas MFN2 has lower GTPase activity and is thought to stabilize the interactions between mitochondria¹⁷. Ablation of *Mfn2* in mice is embryonically lethal and *MFN2* mutations in humans cause Charcot–Marie–Tooth type IIa (CMTIIa), a peripheral sensorimotor neuropathy^{15,18}. Given their role in membrane tethering, MFNs and MFN2 in particular are candidates to modulate ER–mitochondria interactions.

Here we show that MFN2 is enriched at contact sites between ER and mitochondria, regulates morphology of the former and directly tethers the two by means of transorganellar homotypic and heterotypic interactions. The distance between ER and mitochondria increases in cells lacking MFN2, impairing mitochondrial Ca^{2+} uptake. These findings provide the molecular nature of the long sought link between the two organelles, prove the Ca^{2+} microdomains theory and suggest a role for ER and ER–mitochondria juxtaposition in the pathogenesis of CMTIIa.

Mitofusin 2 and ER morphology

We analysed by transmission electron microscopy the effect of *Mfn2* ablation on organelle shape. Together with fragmented mitochondria, several dilated single membrane bound organelles were apparent in *Mfn2*^{−/−} mouse embryonic fibroblasts (MEFs; Supplementary Fig. 1a), which were glucose-6-phosphatase positive and therefore ER in nature (Supplementary Fig. 1b)¹⁹. Dynamin-related proteins modulate the shape of several intracellular organelles^{11,20}. We therefore investigated the role of *Mfn2* in ER morphology, by three-dimensional (3D) reconstruction and volume rendering of confocal stacks of ER-targeted yellow fluorescent protein (ER-YFP). In wild-type and *Mfn1*^{−/−} MEFs, ER appeared as an interconnected network of cisternae spanning the whole cellular volume, which were conversely swollen and aggregated in *Mfn2*^{−/−} cells (Fig. 1a, b). Immunofluorescence analysis of the resident ER protein calreticulin further confirmed ER derangement (data not shown). The altered morphology of *Mfn2*^{−/−} ER resulted from reduced luminal continuity, as shown by fluorescence recovery after photobleaching (FRAP) of ER-YFP (Fig. 1c, d). Furthermore, in HeLa cells the silencing of *MFN2*, but not of *MFN1*, yielded the same mitochondrial and ER phenotypes observed in *Mfn2*^{−/−} MEFs (Supplementary Fig. 2). Re-expression of MFN2, but not of MFN1, corrected ER morphology in *Mfn2*^{−/−} MEFs (Fig. 1a, b). In CMTIIa, mutations cluster in the GTPase domain and in the p21-Ras-binding region that is immediately upstream^{18,21}. MFN2 mutants carrying a point mutation in the GTPase domain (MFN2^{R94Q}) or lacking the p21-Ras-binding domain (amino acids 77–91, MFN2^{ΔRAS}) were expressed as efficiently as wild-type MFN2 (Supplementary Fig. 3), and completely corrected mitochondrial shape (Supplementary Fig. 4) but did not complement ER morphology (Fig. 1a, b).

We were surprised to see ER changes caused by the lack of a mitochondrial protein such as MFN2. Immunofluorescence analysis of endogenous MFN2 in MEFs (data not shown) and HeLa cells (Fig. 2f) showed a reticular staining suggestive of ER localization, in addition to the mitochondrial one. Myc-tagged MFN2 (Myc-MFN2) re-expressed in *Mfn2*^{−/−} cells yielded a similar distribution, with a diffuse mitochondrial (and extra-mitochondrial) signal. Often, spots of intense Myc-MFN2 staining did not overlap with the outer membrane marker

¹Dulbecco–Telethon Institute, Venetian Institute of Molecular Medicine, Via Orus 2, 35129 Padova, Italy. ²Department of Cell Physiology and Metabolism, University of Geneva Medical School, 1 Rue M. Servet, 1211 Geneva, Switzerland.

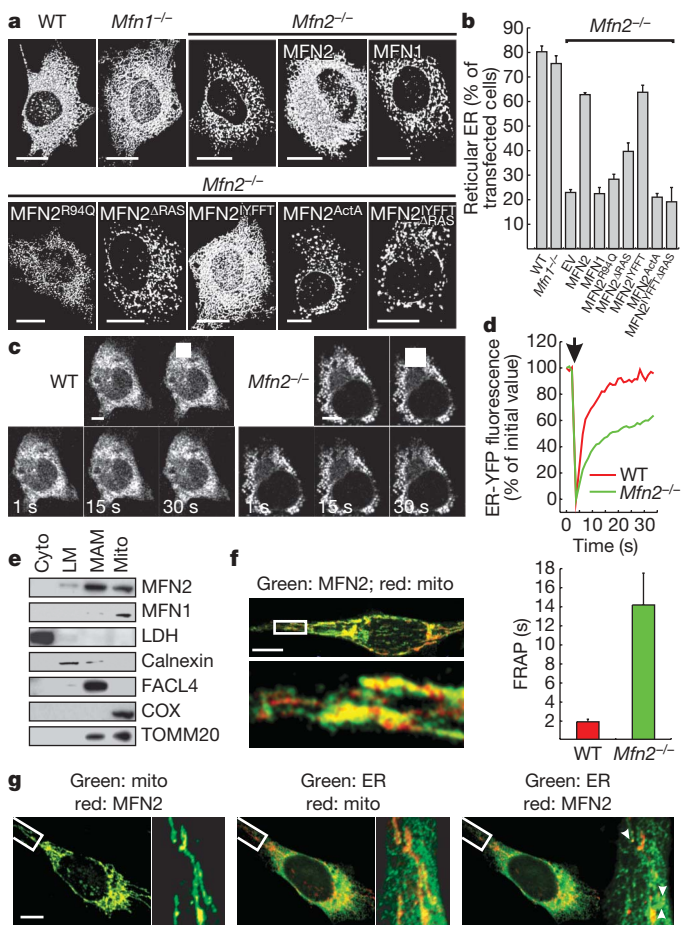


Figure 1 | MFN2 regulates morphology of the endoplasmic reticulum. **a**, Three-dimensional reconstructions of ER in MEFs of indicated genotype. Plasmids co-transfected with ER-YFP are specified. Scale bars, 20 μ m. **b**, Mean and s.e.m. ($n = 8$) of morphometric data from **a**. **c**, Frames from a real-time sequence of ER-YFP FRAP in MEFs. Boxes show photobleached areas. Times (seconds) after photobleaching are indicated. Scale bars, 4 μ m. **d**, Top: ER-YFP fluorescence recordings from experiments as in **c**. The arrow represents photobleaching. Bottom: mean and s.e.m. ($n = 3$, 20 cells per experiment) of data from **c**. **e**, Western blot of the indicated proteins in subcellular fractions from mouse hepatocytes. Cyto, cytosolic fraction; LM, light membranes; mito, mitochondrial fraction. **f**, **g**, Immunofluorescence of MFN2 subcellular localization. HeLa cells were immunostained for TOMM20 (mito) and MFN2 (**f**), and for TOMM20 (mito), MFN2 and calnexin (ER) (**g**). Arrowheads denote spots of MFN2 colocalizing with ER-mitochondria contact points. The boxed areas were magnified $\times 7.4$ (**f**) and $\times 2.7$ (**g**). Scale bars, 10 μ m.

TOMM20 (Supplementary Fig. 5, arrowheads). Endogenous MFN2 also displayed a patchy staining in discrete areas of the outer mitochondrial membrane, overlapping with sites of mitochondria-ER juxtaposition (Fig. 2g). Percoll-purified subcellular fractions devoid of cross-contamination, as indicated by electron microscopy (Supplementary Fig. 6), further substantiated this pattern of subcellular MFN2 distribution. A small amount of MFN2 (6.9%) was retrieved in the ER, whereas most (58.4% of total) was detected in the MAMs fraction, positive for the specific marker FACIL4 (also known as ACSL4; Fig. 2e). Quantitative analysis indicated that MFN2 was enriched 14-fold in MAMs compared to mitochondria (see Supplementary information). Thus, MFN2 is also located at the ER and is enriched at sites of juxtaposition with mitochondria. ER of *Mfn2*^{-/-} cells is fragmented and mutants of MFN2 associated with CMTIIa are unable to correct this ER defect.

ER-mitochondria tethering in *Mfn2*^{-/-} cells

The enrichment of MFN2 in MAMs prompted us to address whether it could participate in ER-mitochondria tethering. Volume-rendered

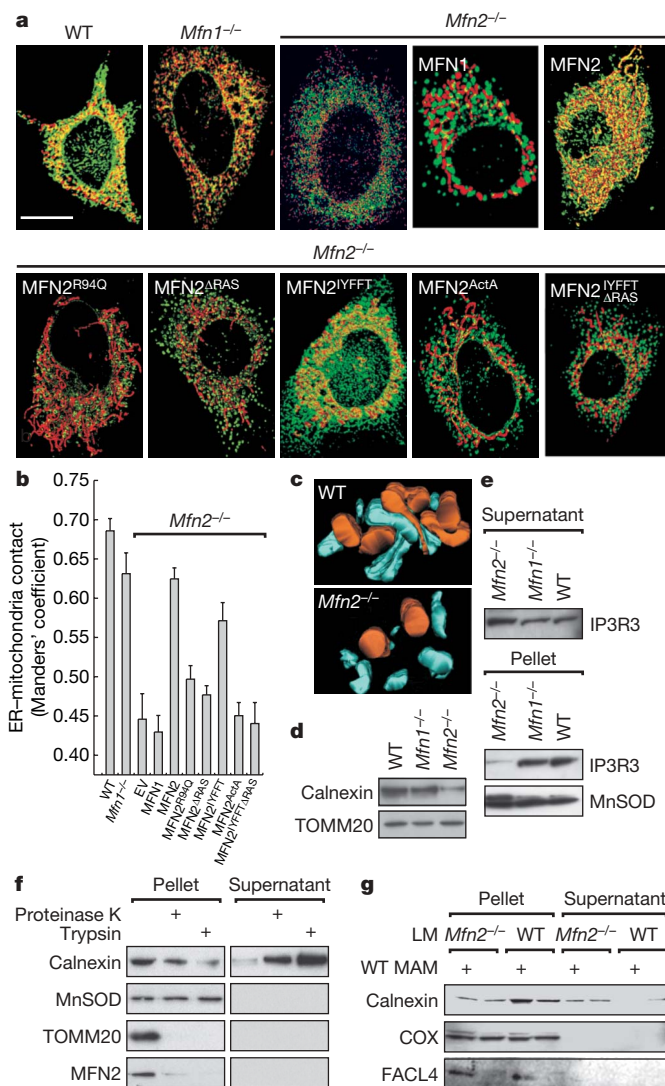


Figure 2 | MFN2 is required at the endoplasmic reticulum to juxtapose it to mitochondria. **a**, Three-dimensional reconstructions of ER and mitochondria in MEFs of indicated genotype. Plasmids co-transfected with ER-YFP and MT-RFP are specified. Yellow indicates that organelles are closer than ~ 270 nm. WT, wild type; scale bar, 18 μ m. **b**, Mean \pm s.e.m. ($n = 4$, 20 cells per experiment) of interaction data from **a**. **c**, Three-dimensional reconstructions of electron tomography of wild-type and *Mfn2*^{-/-} MEFs. Orange denotes mitochondria; cyan represents ER. Scale bar, 500 nm. **d**, Western blot of the specified proteins in crude mitochondrial fractions isolated from the indicated MEFs. **e-g**, *In vitro* assay of ER-mitochondria interaction. Western blots of the indicated proteins in pellets and supernatant after co-sedimentation of mitochondria and microsomes. **e**, Organelles were Percoll-purified from MEFs of the specified genotype. **f**, Organelles were purified from mouse liver. Where indicated, mitochondria were pretreated with proteases to cleave outer membrane proteins. **g**, Mitochondria were isolated from mouse liver and light membranes (LM) from MEFs of the indicated genotype. Where indicated, MAMs from mouse liver were added.

3D reconstructions of z-axis stacks of confocal images of mitochondria- and ER-targeted spectral variants of green fluorescent protein (GFP) has previously been used to measure juxtaposition between the two organelles⁵. In such reconstructions of wild-type MEFs co-expressing ER-YFP and mitochondrial red fluorescent protein (MT-RFP), several yellow areas appeared. In pseudocoloured images this usually indicates 'colocalization'; here, yellow represents areas of organelle tethering below ~ 270 nm (the calculated confocal resolution power) (Fig. 2a). Rotation of the reconstruction on the y -axis corroborated that areas of overlap represent juxtaposition of

organelles (Supplementary Movie 1). Electron tomography allows precise discrimination of the spatial relationship between organelles²² and confirmed close juxtaposition between ER and mitochondria in wild-type MEFs (Fig. 2c and Supplementary Movie 2). Areas of tethering were reduced in *Mfn2*^{-/-}, but not in *Mfn1*^{-/-} MEFs (Fig. 2a), further confirmed by rotation on the y -axis of the confocal reconstruction of *Mfn2*^{-/-} MEFs (Supplementary Movie 3) and by electron tomography (Fig. 2c and Supplementary Movie 4). The reliable Manders' colocalization coefficient²³ revealed a ~40% reduction of the ER–mitochondria juxtaposition in *Mfn2*^{-/-} MEFs (Fig. 2b) and allowed quantitative analysis of complementation by wild-type and mutant MFN2. MFN2—but not MFN1 or the R94Q and Δ RAS mutants—completely corrected tethering in *Mfn2*^{-/-} MEFs (Fig. 2a, b). Genetic and pharmacological analysis excluded that this depended on Ras-regulated cascades (O.M.d.B. and L.S., unpublished observations). MFN2 also proved to be critical for ER–mitochondria juxtaposition in other cell types, as substantiated by silencing of *MFN1* and *MFN2* in HeLa cells (Supplementary Fig. 2). Furthermore, crude *Mfn2*^{-/-} mitochondria isolated by standard differential centrifugation were much less contaminated by microsomes than wild-type or *Mfn1*^{-/-} ones (Fig. 3d).

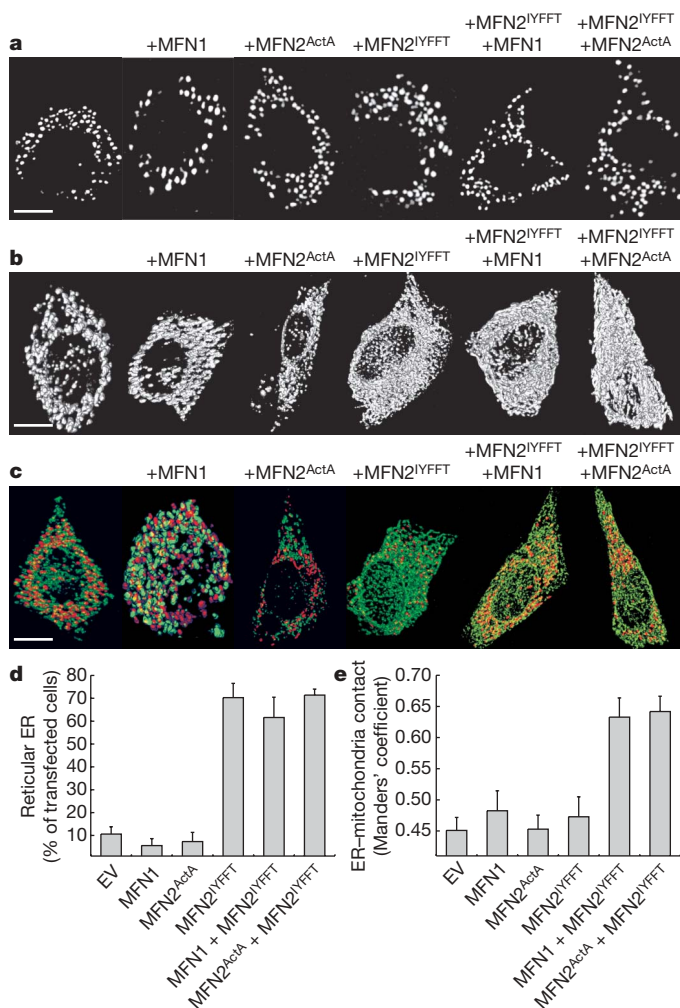


Figure 3 | Juxtaposition between ER and mitochondria requires ER MFN2 and mitochondrial MFN2 or MFN1. **a–c**, *Mfn1*^{-/-} *Mfn2*^{-/-} cells were co-transfected with MT-RFP (**a**), ER-YFP (**b**), or with both (**c**) plus the indicated plasmids, and confocal images of mitochondria (**a**), 3D reconstructions of ER (**b**) or 3D reconstructions of ER and mitochondria (**c**) were acquired. Scale bars, 18 μ m. **d**, Mean and s.e.m. ($n = 7$, 50 cells per experiment) of morphometric data from **a**. **e**, Mean and s.e.m. ($n = 4$, 20 cells experiment) of interaction data from **c**.

ER MFN 2 regulates ER shape and tethering to mitochondria

Selective correction of *Mfn2*^{-/-} MEFs with mutants of MFN2 exclusively targeted to mitochondria or to the ER can highlight its organelle-specific functions. Using established approaches, we constructed mutants of MFN2 whose expression was restricted to the surface of mitochondria (MFN2^{ActA}; ref. 24), or of the ER (MFN2^{cb5}; ref. 24; and MFN2^{YFFT}; ref. 25) (Supplementary Fig. 3). In *Mfn2*^{-/-} cells, mitochondrial elongation induced by MFN2^{ActA} was superimposable to that of MFN2 (Supplementary Fig. 4), whereas ER morphology (Fig. 1a, b) and ER–mitochondria tethering (Fig. 2a, b) were not corrected. MFN2^{YFFT} did not correct mitochondrial morphology (Supplementary Fig. 4), but completely complemented the ER defect (Fig. 1a, b) and restored ER–mitochondria juxtaposition (Fig. 2a, b). Similarly, MFN2^{cb5} was found exclusively in the ER of *Mfn2*^{-/-} MEFs and corrected ER morphology and ER–mitochondria tethering as efficiently as MFN2^{YFFT}, whereas it had no effect on mitochondrial morphology (Supplementary Fig. 7). When the p21-Ras-binding domain was deleted, MFN2^{YFFT} lost its ability to complement ER morphology (Fig. 1a, b) and juxtaposition to mitochondria (Fig. 2a, b).

We extended the analysis of the role of ER MFN2 in tethering ER to mitochondria by devising a reconstituted *in vitro* interaction assay. Highly purified microsomal and mitochondrial fractions (Fig. 1e and Supplementary Fig. 6) were co-incubated and centrifuged at low speeds unable to pellet microsomes (Supplementary Fig. 8). Immunoblotting for an ER-specific marker showed co-sedimentation of microsomes with mitochondria (Fig. 2e). Limited protease digestion confirmed that the interaction between ER and mitochondria requires outer mitochondrial membrane proteins (Fig. 2f). Furthermore, co-sedimentation was inhibited at low temperatures and by a non-hydrolysable analogue of GTP (Supplementary Fig. 8), again suggesting a role for proteins with a GTPase activity (such as MFN2). We therefore compared the efficiency of the interaction in samples from wild-type, *Mfn1*^{-/-} and *Mfn2*^{-/-} MEFs. Whereas wild-type and *Mfn1*^{-/-} microsomes co-sedimented efficiently with genotype-matched mitochondria, between *Mfn2*^{-/-} ER and mitochondria we observed only a residual and unspecific interaction (Fig. 2e), insensitive to boiling of either of the co-incubated fractions (Supplementary Fig. 8). The lack of specific interaction between *Mfn2*^{-/-} ER and mitochondria could suggest a role for ER MFN2 in mediating this interaction. We verified this possibility by comparing the ability of wild-type and *Mfn2*^{-/-} microsomes to interact with wild-type (MFN1⁺, MFN2⁺) mitochondria. Notably, *Mfn2*^{-/-} microsomes were unable to interact above background not only with genotype matched (Fig. 2e), but also with wild-type mitochondria (Fig. 2g). Wild-type MAMs added to the reactions maximized co-sedimentation of wild-type microsomes and mitochondria, but had no effect on the interaction between wild-type mitochondria and *Mfn2*^{-/-} ER (Fig. 3h). Thus, microsomes devoid of MFN2 are unable to engage in productive interactions with wild-type mitochondria even in the presence of MAMs (enriched in MFN2). Thus, the interaction between ER and mitochondria requires MFN2 on the surface of the ER.

ER MFN2 complexes with mitochondrial MFNs

To test the possibility of heterotypic interactions between MFN2 on the ER and MFN1 or MFN2 on mitochondria we used MEFs doubly deficient for *Mfn1* and *Mfn2* (*Mfn1*^{-/-} *Mfn2*^{-/-})²⁶. Ablation of both MFNs caused severe mitochondrial and ER fragmentation, accompanied by a decrease in tethering comparable to that of *Mfn2*^{-/-} cells. As expected, concurrent re-expression of wild-type MFNs completely corrected all the phenotypes (data not shown). MFN1 or MFN2^{ActA} alone did not ameliorate any morphological parameter. Conversely, MFN2^{YFFT} alone complemented ER, but not mitochondrial shape or juxtaposition of the two organelles, confirming that ER–mitochondria juxtaposition depends on factors other than shape of the former. Co-expression of MFN2^{YFFT} with MFN1 or with MFN2^{ActA} fully restored ER–mitochondria tethering (Fig. 3). Thus, MFN2 targeted to the ER

requires a homo- or hetero-partner on the surface of mitochondria to interact *in trans*. Accordingly, haemagglutinin (HA)-tagged MFN1 as well as V5-tagged MFN2^{ActA} co-immunoprecipitated with Myc-MFN2^{YFFFT}, but not with Myc-MFN2^{YFFFTΔRAS} (Fig. 4a) that does not restore ER-mitochondria interaction (Fig. 2a). To verify whether MFN2 on the ER interacts with mitochondrial MFN1 or MFN2 *in situ*, we stabilized protein-protein interaction in live cells with formaldehyde or with disuccinimidyl suberate (DSS), which are membrane permeant, short arm (2 Å for formaldehyde, 11.4 Å for DSS) primary amines cross-linkers. Mitochondrial MFN2^{ActA} or MFN1 and were retrieved in high molecular weight complexes with ER MFN2^{YFFFT} (Fig. 4b), but not MFN2^{YFFFTΔRAS} (Supplementary Fig. 9c, d), distinct from the ones formed by the single mitofusins at 270–500 kDa (Supplementary Fig. 9b and ref. 27). The high molecular weight of these complexes suggests either that we stabilized a multimeric structure of ER and mitochondrial MFNs, or that other proteins participate in these complexes. It is conceivable that these high molecular weight complexes comprise only MFNs on MAM. Indeed, HA-MFN1 that poorly localizes to MAM (Fig. 1e) is less efficiently cross-linked in the complex than MFN2 (Fig. 4b, 19–30% of HA-MFN1 is in the complex). Furthermore, expression of exogenous MFNs could increase the recruitment of MFN molecules in the contact points, enhancing the formation of complexes and the tethering—although the reintroduction of ER MFN2^{YFFFT} in *Mfn2*^{-/-} cells did not enhance juxtaposition above wild-type levels (Fig. 2b). Our genetic and biochemical analysis shows that MFN2 on the ER requires mitochondrial MFN2 or MFN1 to bridge the two organelles.

Mitochondrial Ca²⁺ uptake in *Mfn2*^{-/-} cells

Juxtaposition between ER and mitochondria has been proposed to be crucial for mitochondrial Ca²⁺ uptake during InsP₃-mediated signalling

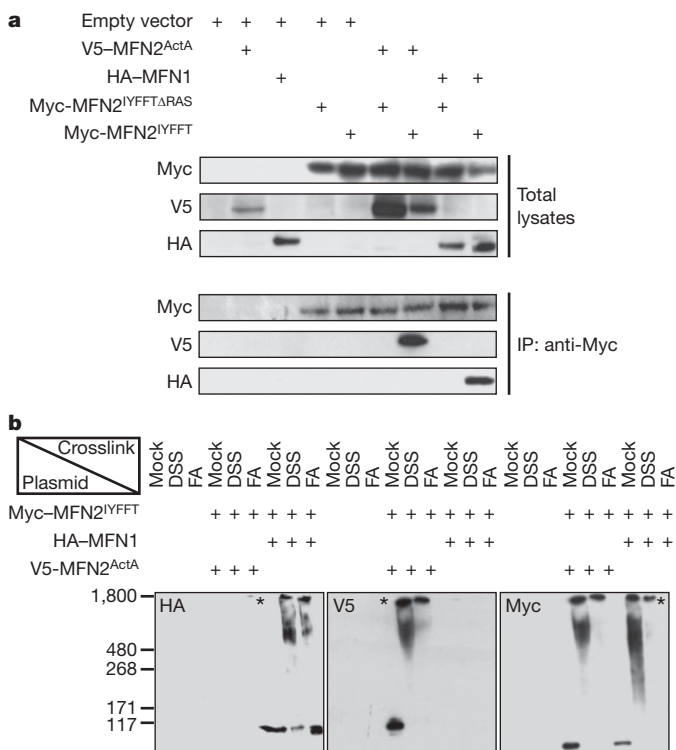


Figure 4 | ER MFN2 engages in homotypic or heterotypic interactions with mitochondrial MFNs. **a**, Western blot of indicated proteins in *Mfn2*^{-/-} transfected with the specified plasmids. IP denotes immunoprecipitation of Myc-MFN2^{YFFFT} using an anti-Myc antibody. **b**, Western blot of indicated proteins in *Mfn2*^{-/-} cells transfected with the specified plasmids and fixed with the indicated cross-linkers. Asterisks indicate a high-molecular weight cross-linked complex containing Myc-MFN2^{YFFFT} and HA-MFN1 or V5-MFN2^{ActA}. FA, formaldehyde.

or passive discharge of the ER Ca²⁺ stores^{4,28}. *Mfn2*^{-/-} cells, in which mitochondria and ER lie distant, are a perfect system to verify this hypothesis. The sarcoplasmic-endoplasmic reticulum Ca²⁺ ATPase (SERCA) inhibitor tert-butyl benzohydroquinone causes the passive release of Ca²⁺ from ER stores and an increase in cytosolic Ca²⁺ concentration ([Ca²⁺]_i). In *Mfn2*^{-/-} MEFs the increase, measured with the fluorescent ratiometric indicator Fura-2, was greater than in wild-type and *Mfn1*^{-/-} cells (Fig. 5a; peak [Ca²⁺]_i ± s.e.m. was 310.3 ± 24.4 nM in wild-type, 297.5 ± 68.1 nM in *Mfn1*^{-/-} and 528.3 ± 67.9 nM in *Mfn2*^{-/-} MEFs; *n* = 8). This increased response of *Mfn2*^{-/-} MEFs was not caused by changes in capacitative calcium entry (Supplementary Fig. 10) and was recapitulated by the InsP₃-coupled agonist ATP (Supplementary Fig. 11). Because [Ca²⁺]_i is modulated by mitochondrial Ca²⁺ uptake, we reasoned that in *Mfn2*^{-/-} MEFs

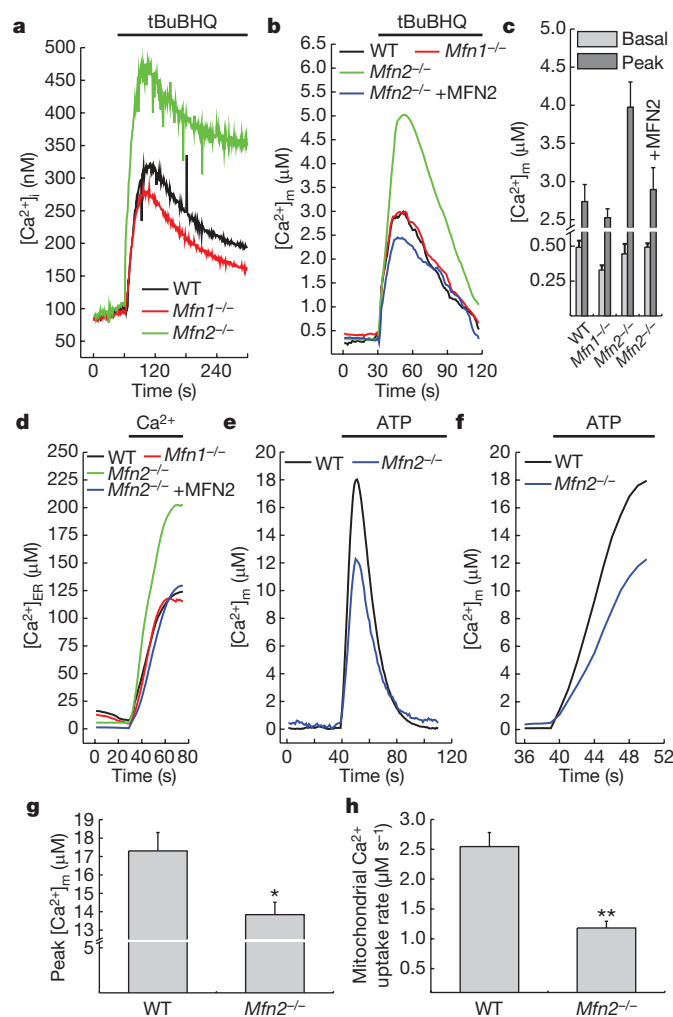


Figure 5 | MFN2 regulates ER Ca²⁺ levels and its transfer to mitochondria.

a, Fura-2 recordings of cytosolic Ca²⁺ ([Ca²⁺]_i) after passive discharge of ER Ca²⁺ stores by tert-butyl benzohydroquinone (tBuBHQ; 0.1 mM) in MEFs of the indicated genotype. **b**, MT-AEQ measurements of [Ca²⁺]_m in response to tert-butyl benzohydroquinone (0.1 mM) in MEFs of the specified genotype. Where specified, MFN2 was co-transfected with ER-AEQ. **c**, Mean and s.e.m. (*n* = 5) of [Ca²⁺]_m from **a**. **d**, ER-AEQ measurements of [Ca²⁺]_{ER} in cells of the indicated genotype. Where specified, MFN2 was co-transfected with ER-AEQ. Stores were replenished by perfusing cells with high Ca²⁺ (1.5 mM) Krebs-Ringer buffer. **e**, MT-AEQ recordings of [Ca²⁺]_m in wild-type and *Mfn2*^{-/-} cells challenged with 0.2 mM or 10 μM ATP, respectively. **f**, Expanded abscissa scale of **e**. **g**, Mean and s.e.m. (*n* = 5) of peak [Ca²⁺]_m from **e**. Asterisk, *P* < 0.05 in a two sample Student's *t* test. **h**, Mean and s.e.m. (*n* = 5) of [Ca²⁺]_m uptake rate from **e**. Double asterisks, *P* < 0.001 in a two sample Student's *t* test.

the decreased ER–mitochondria juxtaposition could cause a reduced mitochondrial Ca^{2+} uptake and therefore the increased $[\text{Ca}^{2+}]_i$ peak. However, in *Mfn2*^{−/−} MEFs, the mitochondrial Ca^{2+} peak in response to discharge of the ER stores—measured using the genetically encoded indicator aequorin—was also higher than in wild-type and *Mfn1*^{−/−} cells. The reintroduction of MFN2 completely corrected this defect (Fig. 5b, c). We therefore turned our attention to the ER: steady state $[\text{Ca}^{2+}]_{\text{ER}}$ was almost doubled in *Mfn2*^{−/−} MEFs (Fig. 5d; peak $[\text{Ca}^{2+}]_{\text{ER}} \pm \text{s.e.m.}$ was $103.8 \pm 21.7 \mu\text{M}$ in wild type, $95.4 \pm 10.5 \mu\text{M}$ in *Mfn1*^{−/−}, $197.3 \pm 29.8 \mu\text{M}$ in *Mfn2*^{−/−} and $93 \pm 13.5 \mu\text{M}$ in *Mfn2*^{−/−} MEFs re-expressing MFN2; $n = 8$), explaining the increased $[\text{Ca}^{2+}]_i$ peak after ER Ca^{2+} release. The increased $[\text{Ca}^{2+}]_{\text{ER}}$ could be sustained by the higher levels of SERCA2 retrieved in *Mfn2*^{−/−} cells (data not shown). Changes in ER morphology and tethering did not depend on this ER Ca^{2+} defect, as they were not corrected in *Mfn2*^{−/−} cells grown in low-extracellular Ca^{2+} buffer that decreases $[\text{Ca}^{2+}]_{\text{ER}}$ ²⁹ (Supplementary Fig. 12). To verify whether the reduced ER–mitochondria juxtaposition affected mitochondrial Ca^{2+} uptake, *Mfn2*^{−/−} mitochondria had to be exposed to InsP_3 -generated Ca^{2+} rises identical to those observed in wild-type MEFs. We titrated the InsP_3 -coupled agonist ATP to generate the same $[\text{Ca}^{2+}]_i$ peak in wild-type and *Mfn2*^{−/−} MEFs (Supplementary Fig. 11). When *Mfn2*^{−/−} mitochondria were exposed to the same $[\text{Ca}^{2+}]_i$ peak, Ca^{2+} uptake was significantly lower and slower (Fig. 5e–h). This difference was not caused by an intrinsic defect of Ca^{2+} handling by *Mfn2*^{−/−} mitochondria, because the rate of mitochondrial Ca^{2+} uptake was superimposable in digitonin-permeabilized wild-type and *Mfn2*^{−/−} MEFs perfused with fixed Ca^{2+} intracellular buffers (data not shown). In conclusion, mitochondrial Ca^{2+} uptake in *Mfn2*^{−/−} MEFs is reduced in response to InsP_3 -generated cytosolic Ca^{2+} transients.

Discussion

We show that MFN2 regulates shape of the ER and tethers it to mitochondria, by engaging in homo- and hetero-complexes comprising MFN2 at the ER and MFN2 or MFN1 on mitochondria. MFN1 and MFN2 are highly homologous, yet they are not functionally redundant. MFN1 is an important molecule for mitochondrial fusion¹³, whereas MFN2 has several extra-mitochondrial functions. Here we extend this to the regulation of ER shape and to its tethering to mitochondria. PACS2 (ref. 10), DRP1 (ref. 11) and the ER protein Nogo-A (also known as RTN4) have all been listed to regulate juxtaposition indirectly, maybe by binding BCL2 on mitochondria³⁰. Conversely, genetic, biochemical and *in vitro* evidence support a direct tethering role for MFN2 on both membranes, particularly on the ER. When MFN2 is ablated, ER coalesces in large vesicles lacking luminal continuity, suggesting a role for MFN2 in ER tubulation. Moreover, MFN2 mutations associated with CMTIIa¹⁸ selectively affect ER tubulation and tethering to mitochondria, pointing to an unexpected pathogenetic mechanism for this disease. Neurodegeneration in CMTIIa could be a consequence of impaired synaptic transmission that requires ER–mitochondria juxtaposition³¹.

ER–mitochondria apposition ultimately affects Ca^{2+} signalling and amplification of apoptotic signals^{2,4,7,32}. Thus, lack of MFN2 has a direct effect on mitochondrial Ca^{2+} uptake during physiological, InsP_3 -mediated release of Ca^{2+} from the ER. High Ca^{2+} microdomains between the latter and mitochondria have been postulated^{2–4} and artificial short staples between the two organelles predispose mitochondria to Ca^{2+} overload²⁸. Along this line, overexpression of MFN2 induces apoptosis³³, perhaps by excessive ER–mitochondria tethering and Ca^{2+} transfer. It would therefore be difficult to assess if increased levels of MFN2 enhance tethering, as it has been performed for the nucleus–vacuole tether Nvj1p in yeast³⁴. Here we provide genetic evidence for the ER–mitochondria Ca^{2+} -microdomains theory: when the distance between the two organelles is increased, such as in *Mfn2*^{−/−} MEFs, Ca^{2+} uptake during InsP_3 -mediated Ca^{2+} signalling is slower.

METHODS SUMMARY

Imaging. For confocal imaging, cells were placed on the stage of a Nikon Eclipse TE300 inverted microscope equipped with a PerkinElmer Ultraview LCI confocal system and a piezoelectric z-axis motorized stage and excited using the appropriate laser lines. For confocal z-axis stacks, 50 images separated by $0.2 \mu\text{m}$ along the z-axis were acquired (total acquisition time = 1.1 s). Three-dimensional reconstruction and volume rendering of the stacks were performed with the appropriate plugins of ImageJ.

Morphometric and contact analysis. Morphometric analysis was performed as described¹³. For ER analysis, cells were scored with reticular ER when major axis was $>5 \mu\text{m}$ and elongation index was >4 for more than 50% of the identified objects. For mitochondria–ER interaction analysis, stacks were automatically thresholded using ImageJ, deconvolved, 3D reconstructed and surface rendered using VolumeJ (ImageJ). Interaction was quantified by Manders' colocalization coefficient²³.

Electron microscopy and tomography. Conventional and tomographic electron microscopy were carried out at the Telethon EM Core Facility as described¹⁴.

Ca^{2+} measurements. Ratiometric cytosolic Ca^{2+} measurements with Fura-2 were determined fluorimetrically as described⁷. Measurements of organellar Ca^{2+} with targeted aequorins were performed as described⁷.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 July; accepted 8 October 2008.

1. Ferri, K. F. & Kroemer, G. Organelle-specific initiation of cell death pathways. *Nature Cell Biol.* **3**, E255–E263 (2001).
2. Rizzuto, R. & Pozzan, T. Microdomains of intracellular Ca^{2+} : molecular determinants and functional consequences. *Physiol. Rev.* **86**, 369–408 (2006).
3. Rizzuto, R., Simpson, A. W., Brini, M. & Pozzan, T. Rapid changes of mitochondrial Ca^{2+} revealed by specifically targeted recombinant aequorin. *Nature* **358**, 325–327 (1992).
4. Rizzuto, R., Brini, M., Murgia, M. & Pozzan, T. Microdomains with high Ca^{2+} close to IP_3 -sensitive channels that are sensed by neighboring mitochondria. *Science* **262**, 744–747 (1993).
5. Rizzuto, R. *et al.* Close contacts with the endoplasmic reticulum as determinants of mitochondrial Ca^{2+} responses. *Science* **280**, 1763–1766 (1998).
6. Hajnoczky, G., Robb-Gaspers, L. D., Seitz, M. B. & Thomas, A. P. Decoding of cytosolic calcium oscillations in the mitochondria. *Cell* **82**, 415–424 (1995).
7. Scorrano, L. *et al.* BAX and BAK regulation of endoplasmic reticulum Ca^{2+} : a control point for apoptosis. *Science* **300**, 135–139 (2003).
8. Vance, J. E. Phospholipid synthesis in a membrane fraction associated with mitochondria. *J. Biol. Chem.* **265**, 7248–7256 (1990).
9. Soltys, B. J. & Gupta, R. S. Interrelationships of endoplasmic reticulum, mitochondria, intermediate filaments, and microtubules—a quadruple fluorescence labeling study. *Biochem. Cell Biol.* **70**, 1174–1186 (1992).
10. Simmen, T. *et al.* PACS-2 controls endoplasmic reticulum–mitochondria communication and Bid-mediated apoptosis. *EMBO J.* **24**, 717–729 (2005).
11. Pitts, K. R., Yoon, Y., Krueger, E. W. & McNiven, M. A. The dynamin-like protein DLP1 is essential for normal distribution and morphology of the endoplasmic reticulum and mitochondria in mammalian cells. *Mol. Biol. Cell* **10**, 4403–4417 (1999).
12. Szabadkai, G. *et al.* Drp-1-dependent division of the mitochondrial network blocks intraorganellar Ca^{2+} waves and protects against Ca^{2+} -mediated apoptosis. *Mol. Cell* **16**, 59–68 (2004).
13. Cipolat, S., de Brito, O. M., Dal Zilio, B. & Scorrano, L. OPA1 requires mitofusin 1 to promote mitochondrial fusion. *Proc. Natl Acad. Sci. USA* **101**, 15927–15932 (2004).
14. Frezza, C. *et al.* OPA1 controls apoptotic cristae remodeling independently from mitochondrial fusion. *Cell* **126**, 177–189 (2006).
15. Chen, H. *et al.* Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *J. Cell Biol.* **160**, 189–200 (2003).
16. Koshiba, T. *et al.* Structural basis of mitochondrial tethering by mitofusin complexes. *Science* **305**, 858–862 (2004).
17. Ishihara, N., Eura, Y. & Mihara, K. Mitofusin 1 and 2 play distinct roles in mitochondrial fusion reactions via GTPase activity. *J. Cell Sci.* **117**, 6535–6546 (2004).
18. Zuchner, S. *et al.* Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot–Marie–Tooth neuropathy type 2A. *Nature Genet.* **36**, 449–451 (2004).
19. Celli, J. *et al.* Brucella evades macrophage killing via VirB-dependent sustained interactions with the endoplasmic reticulum. *J. Exp. Med.* **198**, 545–556 (2003).
20. Koch, A., Yoon, Y., Bonekamp, N. A., McNiven, M. A. & Schrader, M. A role for Fis1 in both mitochondrial and peroxisomal fission in mammalian cells. *Mol. Biol. Cell* **16**, 5077–5086 (2005).
21. Chen, K. H. *et al.* Dysregulation of HSG triggers vascular proliferative disorders. *Nature Cell Biol.* **6**, 872–883 (2004).
22. Marsh, B. J., Mastronarde, D. N., Buttle, K. F., Howell, K. E. & McIntosh, J. R. Organellar relationships in the Golgi region of the pancreatic beta cell line, HIT-

- T15, visualized by high resolution electron tomography. *Proc. Natl Acad. Sci. USA* **98**, 2399–2406 (2001).
23. Manders, E. M., Verbeek, F. J. & Aten, J. A. Measurement of co-localisation of objects in dual-colour confocal images. *J. Microsc.* **169**, 375–382 (1993).
 24. Annis, M. G. *et al.* Endoplasmic reticulum localized Bcl-2 prevents apoptosis when redistribution of cytochrome c is a late event. *Oncogene* **20**, 1939–1952 (2001).
 25. Rojo, M., Legros, F., Chateau, D. & Lombes, A. Membrane topology and mitochondrial targeting of mitofusins, ubiquitous mammalian homologs of the transmembrane GTPase Fzo. *J. Cell Sci.* **115**, 1663–1674 (2002).
 26. Chen, H., Chomyn, A. & Chan, D. C. Disruption of fusion results in mitochondrial heterogeneity and dysfunction. *J. Biol. Chem.* **280**, 26185–26192 (2005).
 27. Karbowski, M., Norris, K. L., Cleland, M. M., Jeong, S. Y. & Youle, R. J. Role of Bax and Bak in mitochondrial morphogenesis. *Nature* **443**, 658–662 (2006).
 28. Csordas, G. *et al.* Structural and functional features and significance of the physical linkage between ER and mitochondria. *J. Cell Biol.* **174**, 915–921 (2006).
 29. Pinton, P. *et al.* The Ca^{2+} concentration of the endoplasmic reticulum is a key determinant of ceramide-induced apoptosis: significance for the molecular mechanism of Bcl-2 action. *EMBO J.* **20**, 2690–2701 (2001).
 30. Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M. & Rapoport, T. A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586 (2006).
 31. Mironov, S. L. & Symonchuk, N. ER vesicles and mitochondria move and communicate at synapses. *J. Cell Sci.* **119**, 4926–4934 (2006).
 32. Ishii, K., Hirose, K. & Iino, M. Ca^{2+} shuttling between endoplasmic reticulum and mitochondria underlying Ca^{2+} oscillations. *EMBO Rep.* **7**, 390–396 (2006).
 33. Guo, X. *et al.* Mitofusin 2 triggers vascular smooth muscle cell apoptosis via mitochondrial death pathway. *Circ. Res.* **101**, 1113–1122 (2007).
 34. Pan, X. *et al.* Nucleus-vacuole junctions in *Saccharomyces cerevisiae* are formed through the direct interaction of Vac8p with Nvj1p. *Mol. Biol. Cell* **11**, 2445–2457 (2000).
 35. Frezza, C., Cipolat, S. & Scorrano, L. Organelle isolation: functional mitochondria from mouse liver, muscle and cultured fibroblasts. *Nature Protocols* **2**, 287–295 (2007).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** O.M.d.B. received a 'Bolsa de Doutoramento' of FCT Portugal. L.S. is senior scientist of the Dulbecco-Telethon Institute and EMBO YIP. This work was supported by Telethon Italy, Compagnia di San Paolo Italy, United Mitochondrial Disease Foundation USA, Muscular Dystrophy Association USA and Swiss National Science Foundation 3100AO-118171.
- Author Contributions** O.M.d.B. and L.S. conceived and designed the experiments and wrote the manuscript. O.M.d.B. performed all the experiments.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.S. (lscorrano@dti.telethon.it).

METHODS

Molecular biology. Details on the plasmid used and on subcloning procedures are available in Supplementary Information.

Cell culture. Wild-type, *Mfn1*^{-/-} and *Mfn2*^{-/-} and *Mfn1*^{-/-} *Mfn2*^{-/-} MEFs were cultured as described before^{15,26} and transfected using Transfectin (Biorad). HeLa cells were transfected with short interfering RNAs (siRNAs) and plasmid DNA using Lipofectamine 2000 (Invitrogen). When indicated, cells were transfected with siRNA 24 h and with MT-RFP/ER-YFP 48 h after seeding.

Imaging. For confocal imaging, cells expressing ER-YFP or MT-RFP were excited using the 488 nm or the 543 nm line of the HeNe laser (PerkinElmer) with exposure times of 50 ms by using a $\times 60$ 1.4 NA Plan Apo objective (Nikon). Confocal z-axis stacks were acquired and 3D reconstruction and volume rendering were performed as described earlier.

Immunofluorescence. Twenty-thousand MEFs of the indicated genotype grown on 13 mm round coverslips, were transfected as indicated and fixed after 24 h. For ER staining, cells were fixed for 15 min at room temperature with 4% ice-cold paraformaldehyde in PBS (pH 7.4), permeabilized for 10 s with ice-cold methanol and incubated for 10 min with 0.38% glycine and 0.27% NH₄Cl in PBS. Cells were blocked by incubation with 3% bovine serum albumin (BSA) in PBS for 30 min and subsequently incubated for 2 h at 25 °C with a rabbit polyclonal anti-Calreticulin antibody (1:500, Upstate).

For anti-MFN2 staining, cells were fixed with 4% paraformaldehyde in PBS, pH 7.4, at room temperature for 30 min, permeabilized with 0.15% Triton X-100 in PBS for 15 min and blocked with 10% BSA in PBS at room temperature for 45 min. Incubation with rabbit anti-MFN2 antibody (1:100 or 1:1,000, gift from R. Youle) and with goat anti-calnexin antibody (1:100, Stressgen) or with mouse anti-TOMM20 antibody (1:100, Santa Cruz) was performed in 10% BSA in PBS at room temperature for 1 h. For anti-Myc immunofluorescence, cells were fixed with 3.7% formaldehyde in PBS, pH 7.4, for 30 min, permeabilized with 0.01% Nonidet P40 for 20 min at room temperature and blocked with 0.5% BSA for 15 min. Incubation with mouse anti-Myc (1:100, Roche) and with rabbit anti-TOMM20 (1:200, Santa Cruz) was performed in PBS at room temperature for 1 h.

Staining of the primary antibody was revealed after incubating for 60 min at room temperature with an anti-rabbit, anti-mouse or anti-goat IgG conjugated to the appropriate fluorochrome. For detection, green channel images were acquired using the appropriate colour channel on the detector assembly of a Nikon Eclipse E600 microscope equipped with a Biorad MRC-1024 laser scanning confocal imaging system with 488 \pm 25 nm bandpass and 525LP filter setting, and with a $\times 63$ 1.4 NA oil immersion objective (Nikon).

Subcellular fractionation. Mouse liver or 10⁹ MEFs of the genotype indicated in the figure legends were washed once with PBS, suspended in isolation buffer (200 mM sucrose, 1 mM EGTA-Tris, and 10 mM Tris-MOPS, pH 7.4), and then disrupted by dounce homogenization. The homogenate was spun at 800g for 10 min; the supernatant was recovered and further centrifuged for 10 min at 8,000g. The resulting pellet (mitochondrial fraction) was collected whereas the

supernatant was further spun for 30 min at 100,000g. The resulting pellet (light membrane fraction) and supernatant (cytosolic fraction) were spun again at 100,000g to further purify the fractions. The mitochondrial fraction was purified further by centrifugation twice at 8,000g for 10 min. The obtained pellet was purified by centrifugation at 95,000g for 30 min on a 30% Percoll gradient in isolation buffer. The mitochondrial layer was washed free of Percoll and resuspended in isolation buffer. Alternatively, subcellular fractions were obtained by differential centrifugation from mouse liver as described³⁵. MAMs were identified as an intermediate layer between the light membranes and the mitochondrial fraction on the Percoll gradient as previously described⁸. Protein concentration was determined using a bichinonate assay (Pierce) and 25 μ g of protein were separated by SDS-PAGE and immunoblotted as indicated in the figure legends.

In vitro ER-mitochondria interaction assay. Subcellular fractions were purified as described earlier and resuspended in reaction buffer (150 mM KCl, 1 mM KH₂PO₄, 5 mM MgCl₂, 10 mM sodium succinate, 10 mM Tris, pH 7.4). Twenty micrograms of mitochondrial protein and 20 μ g of light membrane protein were incubated for 5 min at 37 °C. Subsequently, the reaction was centrifuged at 5,000 g for 5 min at 4 °C. The pellet contained mitochondria and the light membrane fraction interacting with mitochondria, whereas the supernatant contained the remaining light membrane fraction. Where indicated, *in vitro* interaction was performed on mitochondria that were previously boiled (10 min at 95 °C). For digestion of mitochondrial outer membrane proteins, 0.1 μ g ml⁻¹ of mitochondrial protein was digested with 100 μ g ml⁻¹ proteinase K for 30 min on ice or with 5 μ g ml⁻¹ of trypsin for 15 min at room temperature. Reactions were inhibited by the addition of 0.25 mM phenylmethylsulphonyl fluoride. Mitochondria were subsequently pelleted and resuspended in reaction buffer.

Immunoprecipitation, cross-linking and immunoblotting. Details of immunoprecipitation, cross-linking, preparation of samples for immunoblotting and antibodies used for western blotting are available in Supplementary Information.

Transmission electron microscopy and electron tomography. Electron microscopy and electron tomography have been performed as previously described¹⁴. Details of these procedures and of glucose-6-phosphatase staining are available in Supplementary Information.

Fluorescence recovery after photobleaching. Cells were plated and transfected with ER-YFP, and after 24 h placed on the stage of a laser scanning microscope (TCS SP5, Leica). Using the LasAF software (Leica), 16 μ m² manually defined regions were bleached on one z-plane for a total of 3 s using 100% laser power of the 488 nm laser line with a $\times 63$, 1.4NA objective. The post-bleaching images were taken at 1-s intervals for a total of 150 s. Intensities of the photobleached regions were measured and normalized to the intensities of the same region before photobleaching (photobleaching effect) and the neighbouring regions at the same time point (autobleaching effect during the recording).

Ca²⁺ measurements. Measurements of cytosolic [Ca²⁺] using the ratiometric Fura-2 indicator and of organellar [Ca²⁺] using the genetic probe aequorin were performed as described⁷. Details are available in Supplementary Information.

Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins

Heymut Omran^{1*}, Daisuke Kobayashi^{2*†}, Heike Olbrich^{1*}, Tatsuya Tsukahara^{3*}, Niki T. Loges¹, Haruo Hagiwara⁴, Qi Zhang^{5†}, Gerard Leblond⁵, Eileen O'Toole⁶, Chikako Hara⁷, Hideaki Mizuno⁷, Hiroyuki Kawano⁷, Manfred Fliegau¹, Toshiki Yagi^{2†}, Sumito Koshida², Atsushi Miyawaki⁷, Hanswalter Zentgraf⁸, Horst Seithe⁹, Richard Reinhardt¹⁰, Yoshinori Watanabe³, Ritsu Kamiya², David R. Mitchell⁵ & Hiroyuki Takeda²

Cilia and flagella are highly conserved organelles that have diverse roles in cell motility and sensing extracellular signals. Motility defects in cilia and flagella often result in primary ciliary dyskinesia. However, the mechanisms underlying cilia formation and function, and in particular the cytoplasmic assembly of dyneins that power ciliary motility, are only poorly understood. Here we report a new gene, *kintoun* (*ktu*), involved in this cytoplasmic process. This gene was first identified in a medaka mutant, and found to be mutated in primary ciliary dyskinesia patients from two affected families as well as in the *pf13* mutant of *Chlamydomonas*. In the absence of Ktu/PF13, both outer and inner dynein arms are missing or defective in the axoneme, leading to a loss of motility. Biochemical and immunohistochemical studies show that Ktu/PF13 is one of the long-sought proteins involved in pre-assembly of dynein arm complexes in the cytoplasm before intraflagellar transport loads them for the ciliary compartment.

Cilia project from the surface of most eukaryotic cells and consist of a microtubule-based axoneme surrounded by a specialized ciliary membrane. Cilia that contain dynein arms are motile and have diverse roles in cell motility, fluid transport and even in patterning the body of vertebrate embryos. A well-known example is left–right specification determined by cilia-generated fluid flow in the embryonic mouse node¹. Because of their critical and diverse functions, cilia have been linked to a variety of diseases² such as polycystic kidney disease (PKD; Online Mendelian Inheritance in Man (OMIM) accession MIM173900)³ and primary ciliary dyskinesia (PCD; OMIM accession MIM242650)⁴.

PCD, a heterogeneous disorder with an incidence of 1 in 20,000 to 60,000, is characterized by recurrent respiratory tract infections, male infertility and randomization of left–right body asymmetry⁴. The combination of PCD and *situs inversus* is also referred to as Kartagener syndrome (MIM244400). The widespread distributions and functions of motile cilia and flagella explain the several phenotypes of PCD. Recent linkage and mutational analyses with PCD patients have so far identified six genes (*DNAH5*, *DNAI1*, *DNAH11*, *OFD1*, *RPGR* and *TXNDC3*) that are responsible for PCD^{5–10}. Four of these six genes encode subunits of axonemal dyneins, highlighting the importance of dynein motors to ciliary motility.

Medaka fish (*Oryzias latipes*) is an emerging model vertebrate system¹¹ with a high quality draft genome¹² and a number of unique mutants^{13,14}. Our recent mutagenesis screen in medaka identified *kintoun* (*ktu*, formerly named *knt*) as a mutant showing organ laterality defects due to altered ciliary motility¹⁵. Here we show that the *ktu* gene

encodes a cytoplasmic protein conserved from ciliated unicellular organisms to higher mammals. Furthermore, we have identified mutations in the homologous gene of two human PCD families and in *pf13* (ref. 16), a *Chlamydomonas* mutant with paralysed flagella. In all affected organisms the cilia develop normally but have complete or partial loss of outer and inner dynein arms (ODAs and IDAs), leading to loss of motility. Biochemical data demonstrate that Ktu is not a dynein subunit, but is instead required in the cytoplasm for the pre-assembly of dynein arm complexes before they become transported to their destined functional sites within the axoneme¹⁷.

The medaka *ktu* mutant is defective in axonemal dynein arms

ktu is a typical randomized left–right mutant leading to *situs inversus* (Fig. 1a, b). Mutant fish lack directional liquid flow in Kupffer's vesicle (Fig. 1c, d)¹⁵—an organ functionally equivalent to the mouse node in terms of left–right specification^{18,19}. Although the number and length of cilia in Kupffer's vesicle seem normal (Fig. 1e, f), their motility is completely lost (Supplementary Movies 1 and 2). *ktu* homozygous fish are viable but develop PKD (Fig. 1m–o and Supplementary Fig. 2a). Furthermore, male mutant fish have impaired sperm motility leading to reduced fertility (Fig. 1k, l and Supplementary Movies 3 and 4). Ultrastructurally, affected cilia and flagella show partial or complete loss of ODAs and IDAs (Fig. 1g–j). This loss of dynein arms is more severe in Kupffer's vesicle cilia than in sperm flagella (see Supplementary Table 1).

Positional cloning identified a candidate gene encoding 588 amino acids, with a single nucleotide change between homozygous *ktu* and

¹Department of Pediatrics and Adolescent Medicine, University Hospital Freiburg Mathildenstraße 1, D-79106 Freiburg, Germany. ²Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. ³Institute of Molecular and Cellular Biosciences, and Graduate Program in Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Tokyo 113-0032, Japan. ⁴Department of Anatomy and Cell Biology, Gunma University Graduate School of Medicine, Maebashi, Gunma 371-8511, Japan. ⁵Department of Cell and Developmental Biology, SUNY Upstate Medical University, Syracuse, New York 13210-1605, USA. ⁶Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347, USA. ⁷Laboratory for Cell Function Dynamics, Advanced Technology Development Group, Brain Science Institute, RIKEN, Wako, Saitama 351-0198, Japan. ⁸Department of Tumor Virology, German Cancer Research Center, D-69120 Heidelberg, Germany. ⁹Klinik für Kinder und Jugendliche, Klinikum Nürnberg Süd, Breslauer Straße 201, 90471 Nürnberg, Germany. ¹⁰Max-Planck-Institut für molekulare Genetik, D-14195 Berlin, Germany. [†]Present addresses: Department of Anatomy and Developmental Biology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto 602-8566, Japan (D.K.); Structural Biology, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan (T.Y.); Department of Neurobiology, Schering Plough Research Institute, Kenilworth, New Jersey 07033, USA (Q.Z.).

*These authors contributed equally to this work.

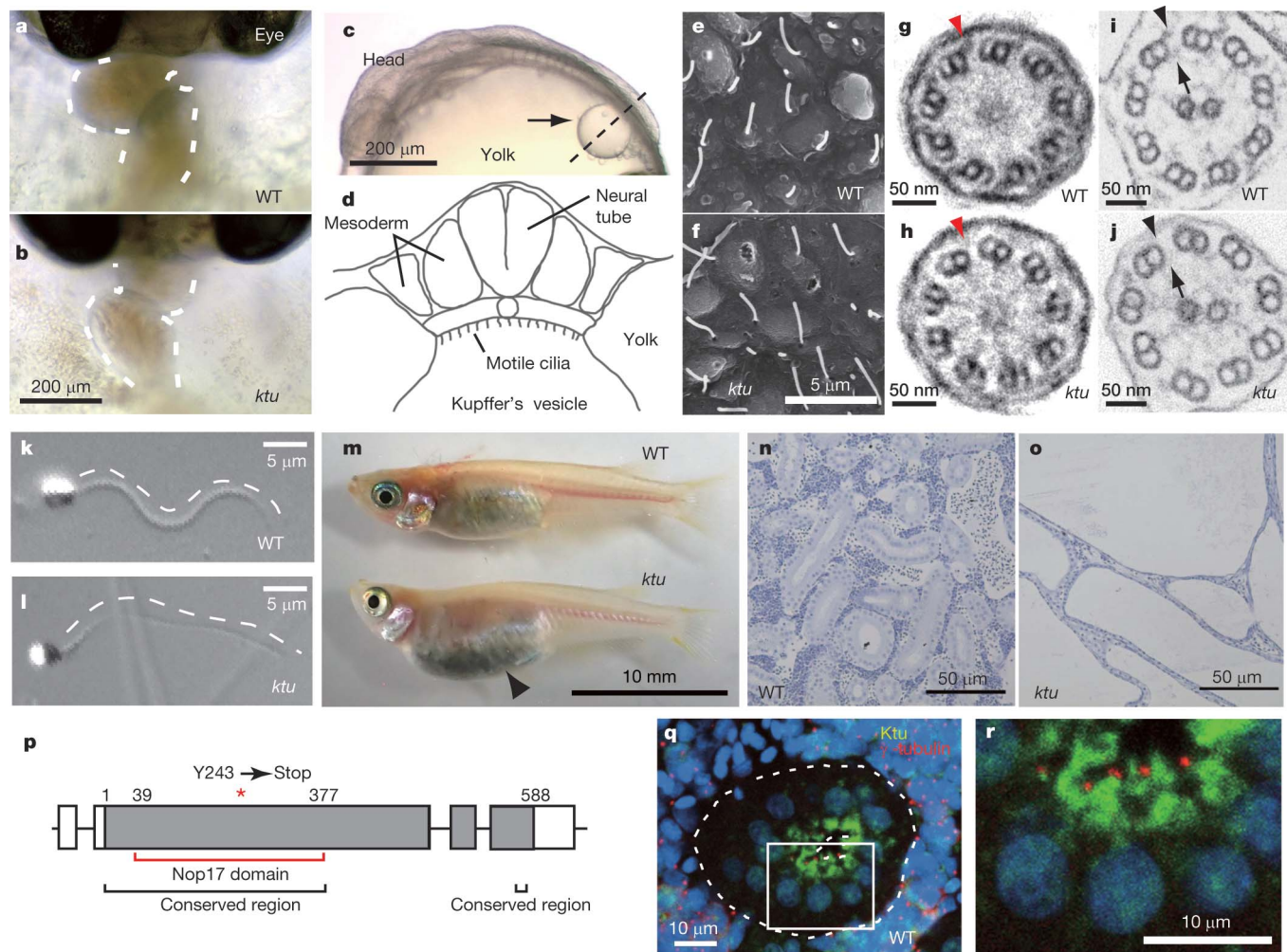


Figure 1 | Medaka *ktu* mutant. **a, b**, Heart looping (dashed line) at 3 days post-fertilization (d.p.f.). **c, d**, Medaka Kupffer's vesicle (arrow in **c**) at 1.5 d.p.f. and its schematic structure. **e, f**, Scanning electron micrographs of cilia in Kupffer's vesicle. **g–j**, Transmission electron micrograph of Kupffer's vesicle cilia (**g, h**; 9+0) and sperm flagella (**i, j**; 9+2), showing a partial or complete loss of ODA (arrowheads) and IDA (arrows) in *ktu*. **k, l**, Swimming sperm. **m**, The belly of the mutant adult (3-month-old) becomes expanded

(arrowhead). **n, o**, PKD in *ktu* mutant kidney. **p**, The exon–intron structure of *ktu*, encoding a protein (solid box) with the conserved N-terminal domain (black bracket) and a domain weakly homologous to yeast Nop17 (red bracket). A premature stop codon introduced in *ktu* is indicated by an asterisk. **q, r**, Overlay and high-power images of a nephric tubule (dotted lines) double-stained with anti-Ktu (green) and γ -tubulin (red) antibodies. Nuclei are in blue.

wild-type embryos, creating a premature stop codon (c.841C > A [p.Y243X]) (in which c denotes complementary DNA, p denotes protein); Fig. 1p and Supplementary Fig. 2d). Knockdown experiments with morpholino antisense oligonucleotides and rescue experiments with wild-type messenger RNA confirmed that we correctly identified the responsible gene (Supplementary Tables 2 and 3). Ktu shows homology to the human hypothetical protein encoded by *C14orf104* (NM_018139.2). Database searches also identified homologous genes in other vertebrates, insects, *Chlamydomonas* and *Tetrahymena* but not in *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae* or *Dictyostelium discoideum* (Supplementary Fig. 2e). The amino-terminal half as well as a short stretch of the carboxy-terminal end is relatively conserved among all Ktu homologues (Fig. 1p and Supplementary Fig. 1). The conserved N-terminal region shows a weak similarity to Nop17 (also known as Pih1), a yeast nucleolar protein involved in assembly of pre-ribosomal-RNA processing complexes^{20–22}.

Polymerase chain reaction with reverse transcription (RT–PCR) analysis detected *ktu* expression in medaka embryos and in nearly all organs of adult mice, with higher expression in mouse tissues known to have motile cilia and flagella, such as brain and testis (data not shown). The subcellular localization of Ktu was then examined using an antibody specific for medaka Ktu. In the stained sperm, the cytoplasm but

not the flagellum was positive for Ktu (Supplementary Fig. 2g). In renal tubules, which are lined by a single layer of tall columnar cells with either single or multiple motile cilia at the apical surface (Supplementary Fig. 2b, c)²³, Ktu is localized in the apical cytoplasm around the γ -tubulin-positive pericentriolar region, not in the cilia (Fig. 1q, r). In most cases, Ktu and γ -tubulin signals did not co-localize. Thus, Ktu is mainly concentrated in the apical cytoplasm.

Mutations in *KTU* cause PCD in humans

The observed *ktu* phenotype in medaka led us to consider the human orthologue *KTU* as a promising candidate gene for human PCD. *KTU* (previously *C14orf104*) is located on chromosome 14q21.3 and consists of three exons encoding cDNAs of 2,511 or 2,367 base pairs (bp; the latter an in-frame splice variant lacking exon 2; Fig. 2c). We have screened affected individuals originating from 112 PCD families for the presence of a mutation in their *KTU* gene. In two consanguineous families we identified homozygous loss-of-function mutations (OP-146: c.1214⁺1215insACGATACCTGCGTGGC [p.G406Rfs89X]; OP-234: c.23C > A [p.S8X]), which is consistent with homozygosity by descent (Fig. 2a, b). All parents were heterozygous for the detected mutations and, in family OP-146, the mutation clearly co-segregated with the disease status. In western blots of cell lysates from control respiratory cells, an anti-KTU antibody detected a specific band at

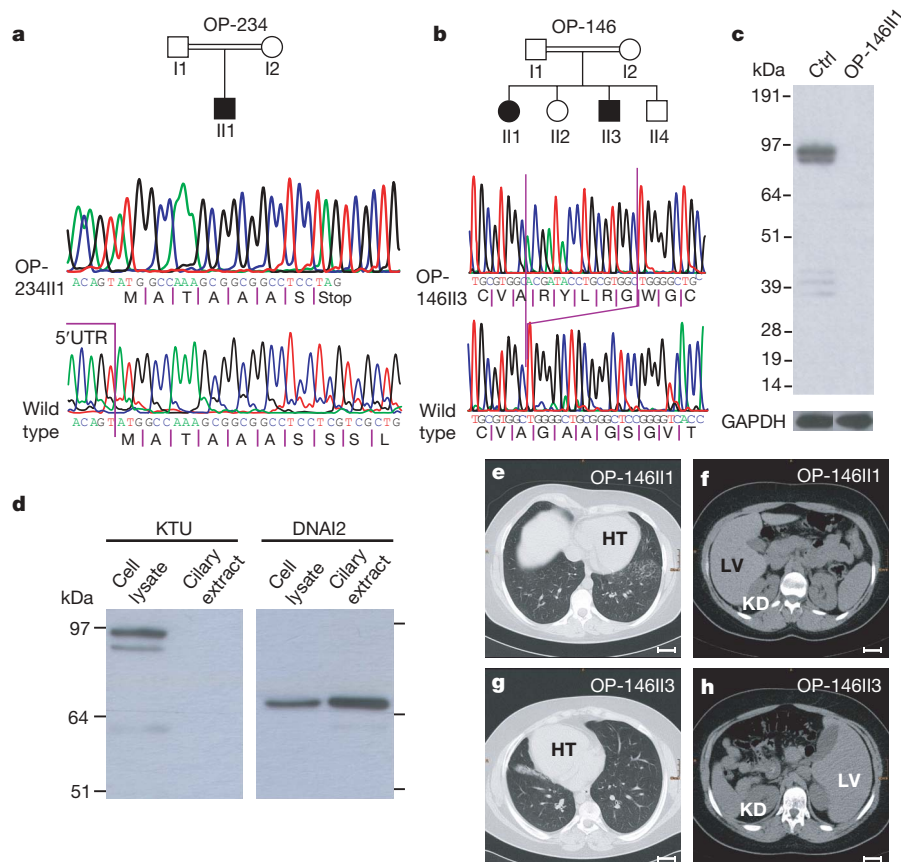


Figure 2 | Recessive loss-of-function *KTU* mutations. **a, b**, Homozygous mutations (c.C23A [pS8X]; c.1214*1215insACGATACCTGCGTGGC [p.G406Rfs89X]) in two PCD patients. **c**, *KTU* is detectable in control respiratory cells (Ctrl) but not in *KTU* mutant cells. The double band is probably explained by distinct alternative splicing products. GAPDH is a loading control. **d**, In control respiratory ciliary extracts *KTU* is absent,

~95 kDa, compatible with the size of the predicted 837-amino-acid *KTU* protein encoded by all three exons (Fig. 2c). In contrast, no specific band was detected for the affected children of family OP-146, confirming a loss-of-function mutation in the *KTU* gene (Fig. 2c). Consistent with the immunohistochemistry in medaka, human *KTU* was detectable in the cytoplasm but not in ciliary fractions of respiratory cells (Fig. 2d).

All three affected patients have suffered from chronic otitis media and sinusitis, and recurrent pneumonia since birth. Two of the affected individuals (OP-146 II3 and OP-234 II1) had developed bronchiectasis. In OP-234 II1, bronchiectasis was so severe that the destroyed middle lobe had to be removed because of persistent pneumonia. Complete *situs inversus* was present in two of the three affected individuals (OP-146 II1 and OP-234 II1) with *KTU* mutations, which is consistent with randomization of left–right body asymmetry (Fig. 2e–h). Notably, their kidneys were of normal size and devoid of cysts (Fig. 2e–h), unlike in medaka *ktu*. This probably reflects a difference in the origin of kidneys: mesonephric for fish and metanephric for mammals. High-speed video microscopy demonstrated complete immotility of respiratory cilia in all three affected patients and of sperm in both affected males (Supplementary Movies 5–10).

Transmission electron microscopy of respiratory cilia and sperm tails identified abnormal axonemal dynein arms (Supplementary Fig. 3a). To further corroborate ultrastructural findings, we performed high-resolution immunofluorescence microscopy in control and *KTU* mutant respiratory cells using specific monoclonal antibodies against dynein arm components (see Supplementary Fig. 4 for antibody specificity and Supplementary Fig. 3 for staining data). We observed the absence of ODA components DNAH5 (γ -heavy chain)

whereas DNAI2 is abundant. *KTU* is only present in the cytoplasm.

e–h, Consistent with randomization of left–right body asymmetry OP-146II1 exhibits *situs solitus* (**e, f**) and her brother (OP-146-II-3) shows *situs inversus totalis* (**g, h**). Both have chronic lung disease. HT, heart; KD, kidney; LV, liver. Scale bars, 3 cm.

and DNAI2 (intermediate chain) from the distal ciliary axonemes, but some residual staining was present in proximal proteins (Supplementary Fig. 3b, d), indicating that loss of *KTU* more strongly affects the assembly of distally located type 2 ODA complexes that specifically contain dynein β -heavy chain orthologue DNAH9 (Supplementary Fig. 3c)²⁴. In sperm, we observed the complete loss of DNAI2 (Supplementary Fig. 5). We then analysed localization of DNALI1—the orthologue of *Chlamydomonas* p28, which is a component of several IDA subtypes^{25,26}. Consistent with IDA defects, no DNALI1 staining was detected in ciliary axonemes of these patients (Supplementary Fig. 3e). Hence in humans, like in medaka, *KTU* mutations cause PCD with combined ODA and IDA defects. Notably, combined ODA and IDA defects were only present in a small proportion of the analysed PCD patients (17 out of 112), of which approximately 12% (2 out of 17) are caused by mutations in the *KTU* gene. In contrast, mutations in *DNAH5*—a main causative gene for PCD²⁷—only affect ODA components without mislocalization of the IDA component DNALI1 (data not shown).

Chlamydomonas PF13 is homologous to vertebrate *ktu*

During this study we found that the *KTU* homologue in *Chlamydomonas* (a single-celled alga used extensively for analysis of cilia and flagella) is PF13, a locus required for ODA assembly¹⁶. The *Chlamydomonas* pf13 gene was identified by cloning DNA from the insertion site of transformation-generated mutant allele pf13-3 (Supplementary Fig. 6)²⁸. Gene identity was confirmed by rescuing the mutant phenotype with wild-type genomic clones (Supplementary Table 5), and by the absence of the gene product in mutant cells (Supplementary Fig. 4f). Sequence of genomic and

cDNA clones revealed that *PF13* encodes a protein homologous to vertebrate Ktu (Supplementary Fig. 1). As expected, western blots show that PF13 is exclusively localized in the cytoplasm but not in flagella (Fig. 3e)—a pattern distinct from that of intraflagellar transport (IFT) components such as IFT46, which are present in both cell body and flagellar fractions (Fig. 3e).

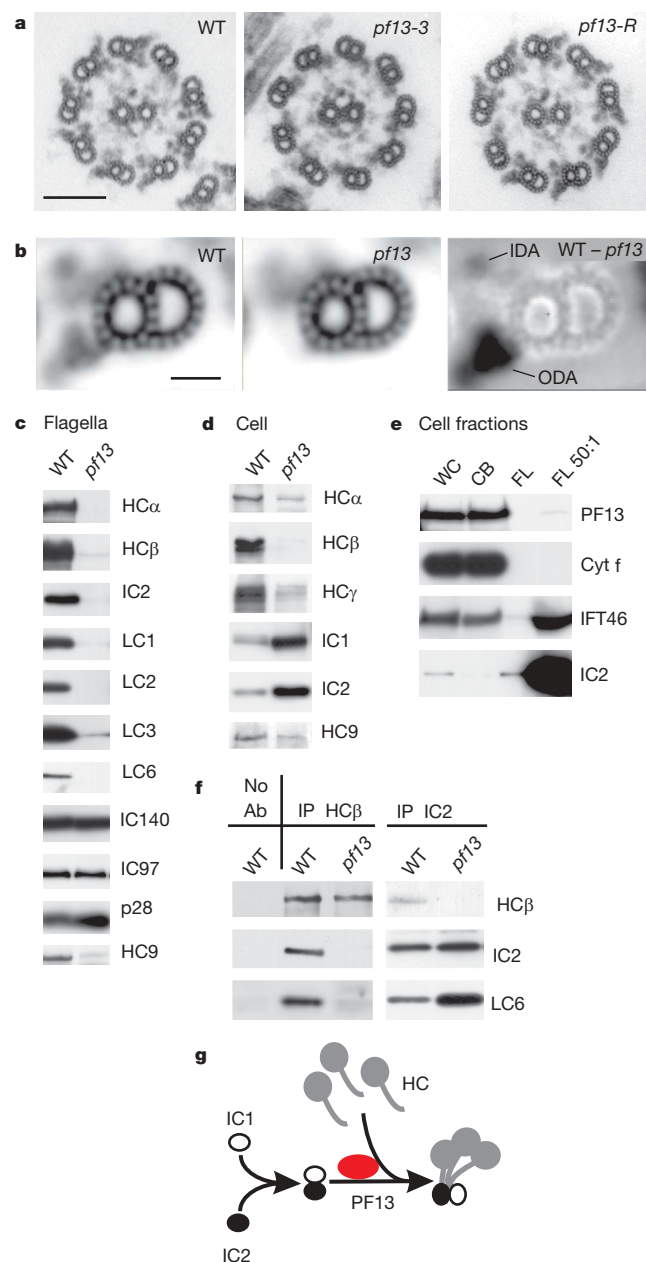


Figure 3 | PF13 is the *Chlamydomonas* homologue of Ktu. **a**, Micrographs of axonemes from wild type (WT), *pf13-3*, and *pf13* transformed with λ 1A (*pf13-R*). Scale bar, 100 nm. **b**, Averages of 108 wild type and 118 *pf13* doublets, and a difference image (WT - *pf13*). Scale bar, 20 nm. **c–f**, Western blots of *Chlamydomonas reinhardtii* proteins. ODA proteins (HC α , HC β , IC2, LC1, LC2 and LC3), and dynein c heavy chain HC9 are depleted from *pf13* flagella but other IDA proteins (IC140, IC97 and p28) are retained (**c**). In *pf13* cytoplasm, ODA and HC9 heavy chains are depleted but ODA ICs are hyper-abundant (**d**). **e**, PF13 occurs in whole cells (WC) and cell bodies (CB) but not in flagella. Absence of cytochrome (Cyt) f demonstrates lack of cell body contamination in flagellar (FL) fractions. FL 50:1, a 50-fold excess of flagella. **f**, Immunoprecipitates of ODA dyneins from wild-type and *pf13* cell extracts probed with antibodies to ODA subunits HC β , IC2 and LC6. **g**, A hypothesized role of PF13 in ODA assembly.

Unlike most other *Chlamydomonas* ODA assembly mutants, which swim slowly with reduced flagellar beat frequencies²⁵, *pf13* mutant cells are paralysed¹⁶. Electron microscopy image averages show that a density is missing from the IDA region in *pf13* axonemes (Fig. 3a, b) consistent with IDA depletion in vertebrate mutant cilia (Fig. 1j and Supplementary Fig. 3e), which probably accounts for the paralysed phenotype²⁵. Western blots show that ODA subunits are greatly depleted from *pf13* flagella, but IDA subunits IC97 and IC140 (needed for assembly of heterodimeric dynein I1) remain at normal levels (Fig. 3c). Unlike in human patients, light chain p28 (a homologue of DNALI1) also remains at wild-type levels, perhaps related to the ability of p28 to assemble independently of dynein heavy chains in *Chlamydomonas*²⁹. Consistently, at least one p28-associated IDA heavy chain, HC9 (the motor subunit of monomeric dynein c), is greatly reduced in *pf13* flagella (Fig. 3c).

Chlamydomonas cells disassemble and reassemble their flagella once each cell cycle, and maintain a sizable cytoplasmic pool of flagellar subunits. Some *Chlamydomonas* dynein assembly mutations prevent pre-assembly of a dynein complex in this cytoplasmic pool¹⁷, whereas others appear to block trafficking of complexes into flagella by IFT³⁰. Western blots show that IDA HC9 and all three ODA heavy chains are depleted from *pf13* cytoplasm, whereas ODA intermediate chains accumulate above normal levels (Fig. 3d), indicating that *pf13* blocks an early step in dynein assembly. To test further the effects of *pf13* on dynein assembly, HC β and IC2 subunits were immunoprecipitated from wild-type and *pf13* cytoplasmic extracts. Western blots (Fig. 3f) confirm that the intermediate chains (IC2) and light chains (LC6) co-precipitate with heavy chains (HC β) from wild-type extracts¹⁷, but show that the small remaining amount of HC β present in *pf13* extracts fails to co-precipitate IC2 or LC6 and thus HC β has not assembled with other subunits. Furthermore, the anti-IC2 antibody co-precipitates IC2 and LC6 (as well as IC1 and LC2; data not shown), demonstrating that intermediate and light chains can still pre-assemble in *pf13* cytoplasm. Thus, PF13 is needed for ODA heavy chain stability, possibly by aiding assembly of heavy chains with other subunits, as diagrammed in Fig. 3g.

Mouse Ktu interacts with Hsp70 and DNAI2

The above results implicate Ktu in the cytoplasmic assembly of dynein complexes in algae. To explore this possibility in vertebrates, we performed immunoprecipitations using rabbit anti-mouse Ktu (mKtu) antibody (Fig. 4a). mKtu is expected to have a similar function to medaka and human counterparts, because mKtu rescued the phenotype of medaka *ktu* mutants (Supplementary Table 4) and *ktu*-knockout mice exhibit PCD-like phenotypes (H.T. and Y. Saga, unpublished data). As shown in Fig. 4b, our immunoprecipitation experiments demonstrated that mKtu co-precipitates with at least one of the intermediate chains, DNAI2 (IC2 in *Chlamydomonas*)—which shows defective assembly in human (Supplementary Fig. 3d) and *Chlamydomonas* (Fig. 3c) mutant cilia. This was further supported by a glutathione S-transferase (GST) pull-down assay using three mKtu constructs (Fig. 4c): both full length and (more weakly) the conserved N-terminal half have the ability to pull down DNAI2, whereas the C-terminal region fails to show such interaction (Fig. 4d).

To obtain mechanistic insight into mKtu function, we looked globally for interacting proteins by applying matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) spectrometry to immunoprecipitates from mouse testis extracts. This analysis identified Hsp70 as one of the main interacting proteins (Supplementary Fig. 7). Among the molecular chaperones examined, Hsp70 and Hsp90 are the most highly expressed in mouse testis (data not shown), but only Hsp70 co-immunoprecipitates with mKtu (Fig. 4b). This was further confirmed by GST-pull-down assay (Fig. 4d).

Discussion

In the present study, forward genetics of medaka and *Chlamydomonas* has been successfully applied to isolate a new gene required in the

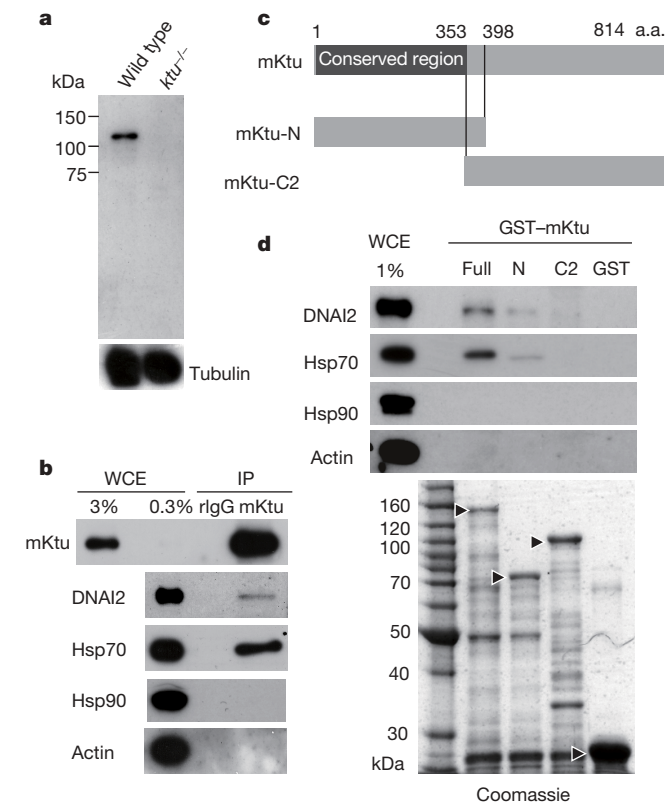


Figure 4 | Ktu binds to dyneins and Hsp70. **a**, The anti-mouse Ktu antibody detects a single band of 110–120 kDa, slightly larger than expected (814 amino acids (aa), 88.3 kDa), which is undetectable in mutant mice. **b**, Immunoprecipitates with anti-mKtu antibody or rabbit (r) IgG (control) from wild-type mouse testis extracts probed with anti-DNAI2, anti-Hsp70, anti-Hsp90 and anti-actin antibodies. **c**, Full and truncated constructs of mKtu used in **d**. **d**, GST pull-down assay. GST-mKtu (full), GST-mKtu-N (N), GST-mKtu-C2 (C2) and GST recombinant proteins (GST) were incubated with mouse testis cytosolic extract. Immunoblots with the indicated antibodies after incubation of GST-recombinant proteins with testis extracts (top). WCE, whole-cell extracts used. The amount of GST-recombinant proteins was confirmed by Coomassie staining (bottom; arrowheads indicate GST-fusion proteins used for the assay).

cytoplasm for ciliary function, which has unique properties compared with other PCD genes. The isolated gene, called *ktu/pf13*, is conserved among animals and ciliated unicellular organisms. As expected, this gene is not found in organisms that lack cilia, such as yeast and *Arabidopsis*, or that have only non-motile cilia (*C. elegans*). Consistently, a previous genomic comparison categorized a partial *Chlamydomonas* PF13 sequence (MOT45), as a gene only found in organisms with motile cilia³¹. Thus Ktu is an ancient protein essential for the motility of all eukaryotic cilia and flagella. The phenotypic similarity of the mutants in three organisms examined suggests that the function of Ktu/PF13 for dynein arm pre-assembly has been evolutionarily conserved from algae to humans.

Our biochemical analyses provided important insight into the proposed function of Ktu/PF13 as a facilitator of dynein pre-assembly, by revealing specific defects in the interaction of intermediate and heavy chains in the cytoplasm (Fig. 3g). Indeed, it has been shown that assembled dynein arms form a cytoplasmic precursor pool and are transferred to a loading zone around the basal body where IFT comes into play^{17,32}. The finding of reduced heavy chain and increased intermediate chain abundance in the *pf13* cytoplasm (Fig. 3d) is consistent with an altered pre-assembly process. Furthermore, immunoprecipitation and GST-pull-down assays demonstrated an interaction between mKtu and one of the intermediate chains, DNAI2. However, only a small fraction of the dynein

co-precipitated with Ktu (Fig. 4), suggesting transient and/or indirect interaction during complex formation. Because Hsp70 was found to be an important interacting protein, Ktu/PF13 could work as a co-chaperone to localize general functions of molecular chaperones (holding, assembling and protecting of proteins from degradation) to dynein arm formation³³. Although Hsp70 is also found in cilia, axonemal Hsp70 is exclusively associated with central pair microtubules, not dyneins³⁴. We therefore believe Hsp70 in these immunoprecipitation fractions is not a subunit of an axonemal dynein, but rather facilitates dynein pre-assembly in conjunction with Ktu. Such a role for Ktu/PF13 would be consistent with the proposed role of its distant homologue, Nop17, as a co-chaperone to Hsp90 for assembly of yeast pre-rRNA processing complexes^{21,22}. Although much progress has been made in understanding the mechanisms of IFT^{35,36}, very little is known about the cytoplasmic process for pre-assembly of dynein complexes. Ktu/PF13 will thus be a good starting point to understand this cytoplasmic process.

METHODS SUMMARY

The medaka *ktu* mutant. *ktu* was isolated in our recent ENU-induced mutagenesis screening¹⁴ as a recessive mutant showing left–right randomization in heart looping.

Antibodies to Ktu and Pf13. We generated polyclonal antibodies to medaka Ktu by immunizing rabbits with two peptides—GSPNPTSDPQENQTRV and GKIKRNERPEHEVKN—to mouse Ktu (with His-tagged full-length Ktu) and to *Chlamydomonas* Pf13 (with GST-fusion protein containing amino acids 3–460). The antibody to human KTU is commercially available, as targeted against anonymous uncharacterized protein C14orf104 (Atlas Antibodies). This antibody did not work well in immunohistochemistry in respiratory cells.

Immunohistochemistry and immunoblotting. For medaka tissues, paraffin-embedded sections (5 μ m) were deparaffined and boiled in 1 mM EDTA for 17 min using a microwave and blocked in 5% skimmed milk, before incubation with primary antibodies. For high-resolution immunofluorescence microscopy and immunoblotting, human respiratory epithelial cells were analysed as previously described²⁴. Monoclonal mouse anti-DNAH5 (amino acids 42–325) and anti-DNAI1 (amino acids 3–280) antibodies were raised by immunization of mice with bacterial-expressed His-tagged protein truncations. Polyclonal rabbit anti-DNAH9 (amino acids 75–220) antibodies were raised by immunization of rabbits with bacterial-expressed maltose-tagged protein truncations.

***Chlamydomonas* pf13 mutant.** The PF13 gene was cloned from the insertional allele *pf13-3* (ref. 28; Supplementary Fig. 6), which was previously created by transformation of a *nit2* strain with plasmid pMN60 and was provided by P. Lefebvre. The *pf13-3* allele was used for all work presented.

Mouse testis cytosolic extract. Seminiferous tubules were taken from the testis of adult male mice (B6). Homogenized samples were filtered through cell strainers and centrifuged. The supernatant was then ultracentrifuged and used as testis cytosolic extracts for immunoprecipitation and pull-down assays.

Patients and mutational analysis. Signed and informed consent was obtained from patients fulfilling diagnostic criteria of PCD⁴. The three *KTU* exons were amplified by PCR and sequenced bidirectionally. In specimens from patients with *KTU* mutations, ciliary and sperm flagella beat was assessed using a high-speed video microscopy system²⁴.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 May; accepted 25 September 2008.

- Okada, Y. *et al.* Mechanism of nodal flow: A conserved symmetry breaking event in left–right axis determination. *Cell* **121**, 633–644 (2005).
- Fliegauf, M., Benzing, T. & Omran, H. When cilia go bad: cilia defects and ciliopathies. *Nature Rev. Mol. Cell Biol.* **8**, 880–893 (2007).
- Wilson, P. D. Polycystic kidney disease. *N. Engl. J. Med.* **350**, 151–164 (2004).
- Zariwala, M. A., Knowles, M. R. & Omran, H. Genetic defects in ciliary structure and function. *Annu. Rev. Physiol.* **69**, 423–450 (2007).
- Olbrich, H. *et al.* Mutations in *DNAH5* cause primary ciliary dyskinesia and randomization of left–right asymmetry. *Nature Genet.* **30**, 143–144 (2002).
- Pennarun, G. *et al.* Loss-of-function mutations in a human gene related to *Chlamydomonas reinhardtii* dynein IC78 result in primary ciliary dyskinesia. *Am. J. Hum. Genet.* **65**, 1508–1519 (1999).
- Bartoloni, L. *et al.* Mutations in the *DNAH11* (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *Proc. Natl Acad. Sci. USA* **99**, 10282–10286 (2002).

8. Budny, B. *et al.* A novel X-linked recessive mental retardation syndrome comprising macrocephaly and ciliary dysfunction is allelic to oral-facial-digital type I syndrome. *Hum. Genet.* **120**, 171–178 (2006).
9. van Dorp, D. B., Wright, A. F., Carothers, A. D. & Bleeker-Wagemakers, E. M. A family with RP3 type of X-linked retinitis pigmentosa: an association with ciliary abnormalities. *Hum. Genet.* **88**, 331–334 (1992).
10. Duriez, B. *et al.* A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proc. Natl Acad. Sci. USA* **104**, 3336–3341 (2007).
11. Wittbrodt, J., Shima, A. & Scharlt, M. Medaka—a model organism from the far East. *Nature Rev. Genet.* **3**, 53–64 (2002).
12. Kasahara, M. *et al.* The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447**, 714–719 (2007).
13. Furutani-Seiki, M. *et al.* A systematic genome-wide screen for mutations affecting organogenesis in Medaka, *Oryzias latipes*. *Mech. Dev.* **121**, 647–658 (2004).
14. Yokoi, H. *et al.* Mutant analyses reveal different functions of *fgfr1* in medaka and zebrafish despite conserved ligand–receptor relationships. *Dev. Biol.* **304**, 326–337 (2007).
15. Hojo, M. *et al.* Right-elevated expression of charon is regulated by fluid flow in medaka Kupffer's vesicle. *Dev. Growth Differ.* **49**, 395–405 (2007).
16. Huang, B., Piperno, G. & Luck, D. J. Paralyzed flagella mutants of *Chlamydomonas reinhardtii* defective for axonemal doublet microtubule arms. *J. Biol. Chem.* **254**, 3091–3099 (1979).
17. Fowkes, M. E. & Mitchell, D. R. The role of preassembled cytoplasmic complexes in assembly of flagellar dynein subunits. *Mol. Biol. Cell* **9**, 2337–2347 (1998).
18. Essner, J. J. *et al.* Kupffer's vesicle is a ciliated organ of asymmetry in the zebrafish embryo that initiates left-right development of the brain, heart and gut. *Development* **132**, 1247–1260 (2005).
19. Kramer-Zucker, A. G. *et al.* Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. *Development* **132**, 1907–1921 (2005).
20. Gonzales, F. A., Zanchin, N. I., Luz, J. S. & Oliveira, C. C. Characterization of *Saccharomyces cerevisiae* Nop17p, a novel Nop58p-interacting protein that is involved in Pre-rRNA processing. *J. Mol. Biol.* **346**, 437–455 (2005).
21. Zhao, R. *et al.* Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J. Cell Biol.* **180**, 563–578 (2008).
22. Boulon, S. *et al.* The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell Biol.* **180**, 579–595 (2008).
23. Mochizuki, E. *et al.* Fish mesonephric model of polycystic kidney disease in medaka (*Oryzias latipes*) pc mutant. *Kidney Int.* **68**, 23–34 (2005).
24. Fliegauf, M. *et al.* Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *Am. J. Respir. Crit. Care Med.* **171**, 1343–1349 (2005).
25. Kamiya, R. Functional diversity of axonemal dyneins as studied in *Chlamydomonas* mutants. *Int. Rev. Cytol.* **219**, 115–155 (2002).
26. LeDizet, M. & Piperno, G. The light chain p28 associates with a subset of inner dynein arm heavy chains in *Chlamydomonas* axonemes. *Mol. Biol. Cell* **6**, 697–711 (1995).
27. Hornef, N. *et al.* DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *Am. J. Respir. Crit. Care Med.* **174**, 120–126 (2006).
28. Tam, L. W. & Lefebvre, P. A. Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. *Genetics* **135**, 375–384 (1993).
29. Yamamoto, R., Yanagisawa, H. A., Yagi, T. & Kamiya, R. A novel subunit of axonemal dynein conserved among lower and higher eukaryotes. *FEBS Lett.* **580**, 6357–6360 (2006).
30. Ahmed, N. T. & Mitchell, D. R. ODA16p, a *Chlamydomonas* flagellar protein needed for dynein assembly. *Mol. Biol. Cell* **16**, 5004–5012 (2005).
31. Merchant, S. S. *et al.* The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245–250 (2007).
32. Ahmed, T. N. *et al.* ODA16 aids axonemal outer row dynein assembly through an interaction with the intraflagellar transport machinery. *J. Cell Biol.* **183**, 313–322 (2008).
33. Young, J. C., Barral, J. M. & Ulrich Hartl, F. More than folding: localized functions of cytosolic chaperones. *Trends Biochem. Sci.* **28**, 541–547 (2003).
34. Mitchell, B. F. *et al.* ATP production in *Chlamydomonas reinhardtii* flagella by glycolytic enzymes. *Mol. Biol. Cell* **16**, 4509–4518 (2005).
35. Hirokawa, N. & Takemura, R. Molecular motors and mechanisms of directional transport in neurons. *Nature Rev. Neurosci.* **6**, 201–214 (2005).
36. Rosenbaum, J. L. & Witman, G. B. Intraflagellar transport. *Nature Rev. Mol. Cell Biol.* **3**, 813–825 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Lo and D. Morris-Rosendahl for critical reading of this manuscript. We are grateful to M. Sugimoto, A. Ito-Igarashi, K. Nakaguchi, S. Minami, Y. H. Park, Y. Mochizuki, Y. Ozawa, K. Ohki, T. Obata, A. Heer and C. Reinhardt for excellent fish care and/or experimental assistance. We also thank A. Shimada and D. Nihei for their help in medaka experiments, J. Freshour and M. Nakatsugawa for help with *Chlamydomonas*, and S. King, H. Qin, W. Sale and D. Stern for antibodies. Our mutant screening was carried out mainly at the National Institute of Genetics (NIG), supported by NIG Cooperative Research Program (2002–2006). This work was supported in part by Grants-in-Aid for Scientific Research Priority Area Genome Science and Scientific Research (A and B), Global COE Program (Integrative Life Science Based on the Study of Biosignaling Mechanisms) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, Yamada Science Foundation, and a Bio-Design Project of the Ministry of Agriculture, Forestry and Fisheries of Japan. D.K. was a research fellow supported by the 21st century COE program of the University of Tokyo, MEXT, Japan. This work was supported by grants to H.Omran from the 'Deutsche Forschungsgemeinschaft' DFG Om 6/4, GRK1104, BIOS and the SFB592, and to D.R.M. from the NIH, GM44228. We would like to acknowledge the sequencing activities by K. Borzym and the Seq-Team at MPI-MG, which was supported by the German Ministry of Science and Education (BMBF) by grant NGFN-2:01GR0414-PDN-S02T17 to R.R. We are grateful for the support by the 'Primare Ciliaere Dyskinesie and Kartagener Syndrom e.V.'.

Author Contributions Research planning and supervision was by H.Omran, D.R.M. and H.T.; medaka genetics and phenotypic analyses by D.K., T.T. and H.T.; biochemical experiments using mouse testis was by T.T., S.K. and Y.W.; high-speed video microscopy of medaka Kupffer's vesicle cilia was by C.H., H.M., H.K., D.K. and A.M.; electron microscopy of medaka cilia/flagella was by H.H. and R.K.; experiments on human PCD were by H. Omran, H. Olbrich, N.T.L., M.F., H.Z., H.S. and R.R.; *Chlamydomonas* experiments were by D.R.M., Q.Z., G.L., E.O., T.Y. and R.K.; and manuscript writing was by H.Omran, D.R.M. and H.T.

Author Information The accession numbers are: medaka *ktu*, AB455535; human *KTU*, FJ158843; mouse *ktu*, AB455811; *Chlamydomonas* PF13 cDNA, AB455237; and *Chlamydomonas* PF13 genome, FJ160770. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.O. (heytmot.omran@uniklinik-freiburg.de) or D.R.M. (MitchellD@upstate.edu) and H.T. (htakeda@biol.s.u-tokyo.ac.jp).

METHODS

Immunohistochemistry. Renal tubules were fixed in 4% paraformaldehyde (PFA) in PBS-Tween20 (PBST) for 1 h. After dehydration, they were paraffin-embedded and sectioned at 5 μ m. The sections were deparaffined and boiled in 1 mM EDTA for 17 min using a microwave and blocked in 5% skimmed milk in PBST. The Ktu signal was amplified using a VECTASTAIN ABC elite kit (Vector Laboratory) and TSA-Alexa647 (Molecular Probes).

Antibodies used include mouse monoclonal antibodies to α -tubulin (DM1A, Sigma), γ -tubulin (GTU-88, Sigma) and acetylated- α -tubulin (6-11B-1, Sigma). Secondary antibodies (Alexa Fluor 488 and Alexa Fluor 555) were from Molecular Probes/Invitrogen. DNA was stained with propidium iodide (Sigma).

Video microscopy of cilia. Medaka sperm motility was captured using $\times 40$ lens on Olympus BX61 equipped with ARTCAM-130MI-BW (ARTRAY). Movies were acquired at about 100 frames per second and played back at 30 frames per second.

Kupffer's vesicle cilia movement was captured using a $\times 60$ water immersion lens on an Olympus high-speed microscope system (Olympus) equipped with FASTCAM-MAX (Photron).

Electron microscopy for medaka. Before fixation, medaka embryos were dechorionated using hatching enzyme in 1 \times Yamamoto Ringer. Dechorionated embryos were incubated in 1 \times Yamamoto ringer until the 8-somite stage. Embryos were fixed and processed as previously described³⁷. For electron microscopy of the sperm flagella, collected sperm were processed as described³⁸.

GST pull-down assay. pGEX 4T-2 plasmids (GE Healthcare) containing mKtu full-length/N/C2 cDNA were used. Transformed and IPTG-induced *Escherichia coli* (BL21 codon+) were sonicated in solubilization buffer (40 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 0.5% Triton X-100). After centrifugation, supernatants were mixed with glutathione sepharose 4B (GE Healthcare) at 4 °C. Glutathione beads were washed with solubilization buffer three times and resuspended with 100 μ l of binding buffer (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 5 mM MgCl₂, 0.8% NP-40).

GST and GST-mKtu full length/N/C2 proteins were conjugated with glutathione beads. After each bead was resuspended with 100 μ l of binding buffer, 200 μ l of mouse testis cytosolic extracts were added and mixtures were rotated at 4 °C for 40 min. Beads were washed with binding buffer three times, proteins were separated by SDS-PAGE and analysed by western blotting using anti-DNAI2 monoclonal antibody (M01, clone 1C8, Abnova), and anti- α -tubulin antibody (DM1A, Sigma). Additional antibodies against Hsp70 (BD Biosciences, 610607), Hsp90 (BD Biosciences, 610418) and actin (polyclonal anti-goat IgG, Santa Cruz) were also used.

Immunoprecipitation from mouse testis extracts. Anti-mKtu antibody and rabbit control IgG were chemically cross-linked to protein A beads using DMP (dimethyl pimelimidate, Sigma). Prepared mouse testis extracts were mixed with 1/4 volume of IP buffer (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 5 mM MgCl₂, 0.8% NP-40), and then ~ 15 μ l of anti-mKtu/rIgG beads (1 μ g μ l⁻¹) was added. The mixture was rotated for 3 h at 4 °C and then washed with ~ 20 -fold reaction-volume (the volume for testis extracts incubated with anti-mKtu/rIgG beads) of IP buffer. Immunoprecipitates were eluted from anti-body-coupled beads by elution buffer (100 mM glycine, pH 2.0, 150 mM NaCl) and neutralized by 1/10 eluted volume of 1.5 M Tris-HCl, pH 8.8.

For mass-spectrometric analysis, see Supplementary Fig. 7.

Immunoblotting. Protein extracts were prepared from human respiratory epithelial cells. Proteins were separated on a NuPAGE 4–12% Bis-Tris gel (Invitrogen) and blotted onto a PVDF membrane (Amersham). The blot was processed for ECL plus (Amersham) detection using anti-KTU (Atlas antibodies) or anti-GAPDH (1:1,000) primary antibodies, and anti-rabbit/anti-mouse HRP secondary antibodies (1:5,000; Santa Cruz).

Immunofluorescence analysis for human respiratory cells. Respiratory epithelial cells were obtained by nasal brush biopsy (cytobrush plus, Medscand Malmö) and suspended in cell culture medium. Sperm cells were washed with PBS. Samples were spread onto glass slides, air dried and stored at -80 °C until use. Staining was performed as described previously²⁴. The antibodies used were: mouse anti-acetylated- α -tubulin antibody (Sigma), monoclonal mouse anti-DNAI2 antibody (Abnova Corporation) and polyclonal rabbit anti- α / β -tubulin antibody (Cell Signaling Technology Inc.). Highly cross-adsorbed secondary antibodies (Alexa Fluor 488, Alexa Fluor 546) were obtained from Molecular Probes (Invitrogen). DNA was stained with Hoechst 33342 (Sigma).

Patients and families. Signed and informed consent was obtained from patients and family members using protocols approved by the Institutional Ethics Review

Board at the University of Freiburg. We studied DNA from a total of 112 PCD patients originating from unrelated families. Clinical information of the three patients carrying homozygous *KTU* mutations was obtained by review of the medical records.

Mutational analysis. Genomic DNA was isolated by standard methods directly from blood samples or from lymphocyte cultures after Epstein–Barr virus transformation. The three exons of *KTU* (*C14orf104*) were amplified by PCR of three genomic fragments and sequenced bidirectionally using BigDye Terminator v3.1 Cycle Sequencing Kit (Perkin Elmer). Samples were analysed on an Applied Biosystems 3730xl DNA Analyser. Sequence data were evaluated using the CodonCode software (CodonCode Corporation). Segregation analyses were performed by sequencing.

High-speed video analysis of ciliary beat in human cells. Ciliary and sperm flagella beat was assessed with the SAVA system (Sisson-Ammons Video Analysis of ciliary beat frequency). Transnasal brush biopsies and sperm ejaculates were rinsed in cell culture medium and immediately viewed with an Olympus IMT-2 inverted phase-contrast microscope equipped with a Redlake ES-310 Turbo monochrome high-speed video camera (Redlake) and a $\times 40$ objective. Digital image sampling was performed at 125 frames per second and 640 \times 480 pixel resolution.

Electron microscopy of *Chlamydomonas*. Axonemes were isolated, fixed and processed for thin section electron microscopy as previously described³⁹. Averages were generated using digital images from scanned negatives and were normalized to a region on the outside wall of the doublet microtubules. Normalization, averaging and creation of subtracted images followed previously described procedures⁴⁰.

Cloning the *PF13* gene. See Supplementary Fig. 6.

Antibodies to *Chlamydomonas* proteins. A portion of the *PF13* cDNA spanning amino acids 3–460 was cloned into pGEX-2T, and a bacterially expressed GST-fusion protein was gel-purified and used to raise two rabbit polyclonal antisera: R1087 and R1088.

Additional antibodies used include polyclonal anti-IFT46 antibody⁴¹, polyclonal anti-HC α antibody¹⁷, monoclonal anti-HC β antibody, anti-IC1 and anti-IC2 antibodies⁴², monoclonal anti-HC γ (12 γ B) antibody⁴³ and polyclonal anti-LC1, -LC2, -LC3 and -LC6 antibodies⁴⁴. Antibodies against inner row dyneins included polyclonal anti-p28 antibody²⁶, polyclonal anti-IC97 antibody (W. Sale), polyclonal anti-IC140 antibody⁴⁵ and polyclonal anti-HC9 antibody. The HC9 antibody was generated by cloning nucleotides 859–1421 of the *HC9* cDNA coding region, spanning amino acids 287–473 (ref. 46), into pGEX2T and using the resulting GST-fusion protein to immunize rabbits (see Supplementary Fig. 4e for antibody specificity). As a marker for cytoplasmic proteins, we used a polyclonal antibody against chloroplast cytochrome f, the product of the chloroplast *PetA* gene⁴⁷. Immunoprecipitation was performed as previously described¹⁷.

37. Hagiwara, H., Shibasaki, S. & Ohwada, N. Abnormal cilia in human uterine tube epithelium. *J. Clin. Electron Microsc.* **23**, 493–503 (1990).
38. Kamiya, R. Mutations at twelve independent loci result in absence of outer dynein arms in *Chlamydomonas reinhardtii*. *J. Cell Biol.* **107**, 2253–2258 (1988).
39. Mitchell, D. R. & Sale, W. S. Characterization of a *Chlamydomonas* insertional mutant that disrupts flagellar central pair microtubule-associated structures. *J. Cell Biol.* **144**, 293–304 (1999).
40. Mastrorade, D. N. et al. Arrangement of inner dynein arms in wild-type and mutant flagella of *Chlamydomonas*. *J. Cell Biol.* **118**, 1145–1162 (1992).
41. Hou, Y. et al. Functional analysis of an individual IFT protein: IFT46 is required for transport of outer dynein arms into flagella. *J. Cell Biol.* **176**, 653–665 (2007).
42. Mitchell, D. R. & Rosenbaum, J. L. Protein–protein interactions in the 18S ATPase of *Chlamydomonas* outer dynein arms. *Cell Motil. Cytoskeleton* **6**, 510–520 (1986).
43. King, S. M., Otter, T. & Witman, G. B. Characterization of monoclonal antibodies against *Chlamydomonas* flagellar dyneins by high-resolution protein blotting. *Proc. Natl Acad. Sci. USA* **82**, 4717–4721 (1985).
44. DiBella, L. M. et al. Differential light chain assembly influences outer arm dynein motor function. *Mol. Biol. Cell* **16**, 5661–5674 (2005).
45. Yang, P. & Sale, W. S. The Mr 140,000 intermediate chain of *Chlamydomonas* flagellar inner arm dynein is a WD-repeat protein implicated in dynein arm anchoring. *Mol. Biol. Cell* **9**, 3335–3349 (1998).
46. Yagi, T. et al. An axonemal dynein particularly important for flagellar movement at high viscosity. Implications from a new *Chlamydomonas* mutant deficient in the dynein heavy chain gene DHC9. *J. Biol. Chem.* **280**, 41412–41420 (2005).
47. Chen, X., Kindle, K. L. & Stern, D. B. The initiation codon determines the efficiency but not the site of translation initiation in *Chlamydomonas* chloroplasts. *Plant Cell* **7**, 1295–1305 (1995).

Tycho Brahe's 1572 supernova as a standard type Ia as revealed by its light-echo spectrum

Oliver Krause¹, Masaomi Tanaka^{2,3}, Tomonori Usuda⁴, Takashi Hattori⁴, Miwa Goto¹, Stephan Birkmann^{1,5} & Ken'ichi Nomoto^{2,3}

Type Ia supernovae are thermonuclear explosions of white dwarf stars in close binary systems¹. They play an important role as cosmological distance indicators and have led to the discovery of the accelerated expansion of the Universe^{2,3}. Among the most important unsolved questions⁴ about supernovae are how the explosion actually proceeds and whether accretion occurs from a companion or by the merging of two white dwarfs. Tycho Brahe's supernova of 1572 (SN 1572) is thought to be one of the best candidates for a type Ia supernova in the Milky Way⁵. The proximity of the SN 1572 remnant has allowed detailed studies, such as the possible identification of the binary companion⁶, and provides a unique opportunity to test theories of the explosion mechanism and the nature of the progenitor. The determination of the hitherto unknown^{7–9} spectroscopic type of this supernova is crucial in relating these results to the diverse population of type Ia supernovae¹⁰. Here we report an optical spectrum of Tycho's supernova near maximum brightness, obtained from a scattered-light echo more than four centuries after the direct light from the explosion swept past the Earth. We find that SN 1572 belongs to the majority class of normal type Ia supernovae.

The supernova of 1572 marked a milestone in the history of science. Danish astronomer Tycho Brahe concluded from his accurate observations of the 'new star' in the constellation of Cassiopeia that it must be located far beyond the Moon¹¹. This contradiction to the Aristotelian concept, that a change on the sky can only occur in the sublunar regime, ultimately led to the abandonment of the notion of the immutability of the heavens. The classification of SN 1572 has

been controversial. On the basis of historical records of the light curve and colour evolution, it has been interpreted as a type Ia supernova of either a normal or somewhat overluminous type⁷ or of a subluminous type⁸. Core-collapse supernovae of type Ib⁹ or II-L¹² have also been suggested to be compatible with the light curve. Support for the interpretation as a type Ia has been inferred from X-ray studies of the ejecta composition¹³, but the determination of the exact supernova type has not been possible without spectroscopic information.

The discovery of light echoes from historic Galactic supernovae, due to both scattering and absorption/re-emission of the outgoing supernova flash by the interstellar dust near the remnant^{14,15}, raised the opportunity of conducting spectroscopic 'post mortems' of historic Galactic supernovae. Such a precise determination of the spectral type long after the original explosion has recently been performed for the Cassiopeia A supernova¹⁶ and a supernova in the Large Magellanic Cloud¹⁷.

We obtained Johnson R-band images of the recently identified¹⁵ light-echo fields of SN 1572 using the 2.2-m and 3.5-m telescopes at the Calar Alto Astronomical Observatory, Spain, on 23 August 2008 and 2 September 2008. One of the fields observed on 2 September 2008, at an angular distance of $d = 3.15^\circ$ and position angle of 62° from the SN 1572 remnant, showed a bright and extended nebula with a peak surface brightness of $R = 23.6 \pm 0.2$ mag arcsec⁻² (Fig. 1a). The position of this new emission feature, relative to the previously reported light-echo detection, has shifted away from the SN 1572 remnant.

The region was re-observed using the Faint Object Camera and Spectrograph (FOCAS) at the Subaru 8.2-m telescope on Mauna Kea,

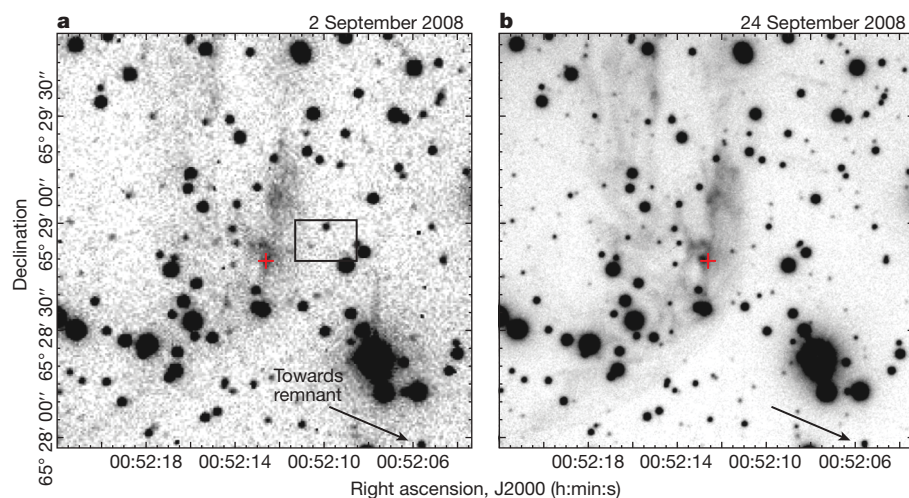


Figure 1 | Optical images of the SN 1572 light echo. **a, b**, R-band images of the same 120×120 -arcsec² area. The corresponding observing epochs are labelled. The position of the brightness peak in the first epoch is marked for reference (red cross). The rectangle shown in **a** indicates the location of a previous light-echo detection¹⁵. The vector towards the remnant of SN 1572 is indicated (arrow). The respective seeing for **a** and **b** were 1.5 and 0.9 arcsec, full-width at half-maximum. The integration times of the two images were 20 and 12 min, respectively. Image reduction was performed using standard methods with the Image Reduction and Analysis Facility software.

¹Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany. ²Institute for the Physics and Mathematics of the Universe, University of Tokyo, Kashiwanoha 5-1-5, Kashiwa, Chiba 277-8568, Japan. ³Department of Astronomy, Graduate School of Science, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan. ⁴SUBARU Telescope, National Astronomical Observatory of Japan, 650 North A'ohoku Place, Hilo, Hawaii, USA. ⁵European Space Agency, Space Science Department, Keplerlaan 1, 2200 AG Noordwijk, The Netherlands.

Hawaii, on 24 September 2008 (Fig. 1b). The peak of the emission, with a surface brightness of $R = 23.5 \pm 0.2$ mag arcsec $^{-2}$, has again shifted away from SN 1572. The shift of 1.4 ± 0.2 arcsec within 22 days is consistent with a light-echo origin. A long-slit spectrum of the brightness peak of the echo structure at position RA 00 h 52 m 12.79 s, dec. $65^\circ 28' 49.7''$ (J2000) was obtained with FOCAS on the same night, covering the wavelength range from 3,800 to 9,200 Å with a spectral resolution of 24 Å.

The acquired echo spectrum unambiguously shows light of a supernova origin (Fig. 2). A number of broad absorption and emission features from neutral and singly ionized intermediate mass elements were detected, all of which are commonly observed in supernovae^{10,18}. Type I supernovae are distinguished from those of type II by the absence of hydrogen, and type Ia supernovae are further distinguished from types Ib and Ic by a prominent silicon 6,355 Å absorption feature at maximum light. This feature is clearly seen in the SN 1572 spectrum as a deep absorption minimum at 6,130 Å with a width of 9,000 km s $^{-1}$ at half maximum. The absorption minimum of the line corresponds to a velocity of 12,000 km s $^{-1}$, typical for normal type Ia supernovae at maximum brightness^{10,18}. Other strong

features detected in the spectrum are Si II 4,135 Å, Fe II, Fe III, Na I D + Si II 5,972 Å, O I 7,774 Å and the Ca II infrared triplet.

The echo spectrum represents supernova light during an interval of time around maximum brightness being averaged over the spatial extent of the scattering cloud. We therefore compared the echo spectrum with the spectra of other type Ia supernovae time-averaged over the brightness peak of the light curve from 0 to 90 days after explosion. The light-echo spectrum of SN 1572 matches such comparison spectra of four well-observed normal type Ia supernovae (1994D, 1996X, 1998bu, 2005cf) and a type Ia composite spectrum¹⁹ remarkably well. Even faint notches observed in normal type Ia spectra at 4,550, 4,650 and 5,150 Å (ref. 18) can be recognized. Values of the reduced chi-squared value from the comparison range between $\chi^2 = 1.5$ and $\chi^2 = 2.5$. The agreement between the spectra indicates that the scattering dust cloud is homogeneous on a length scale of at least 90 light days.

We have compared the spectrum of SN 1572 with thermonuclear supernovae of different luminosity. Both sub- and overluminous type Ia supernovae, such as SN 1991bg²⁰ and SN 1991T²¹, respectively, showed peculiarities in their spectra near maximum light. SN 1991T lacked a well-defined Si II 6,355 Å absorption feature at maximum light, although the subsequent evolution was similar to normal type Ia supernovae. The lack of Si II absorption is visible as an imprint in the time-averaged spectrum and differs from the strong Si II feature in our spectrum of SN 1572. The class of subluminoous objects shows a characteristic deep absorption trough at a wavelength of 4,200 Å, attributed to Ti II (ref. 20), near maximum light. Such a feature is not seen for SN 1572. The sub- and overluminous type Ia templates, obtained in the same way as described in ref. 19 (template spectra available at http://supernova.lbl.gov/~nugent/nugent_templates.html), do not provide a good match to our spectrum of SN 1572, with respective values of $\chi^2 = 8.6$ and $\chi^2 = 10.1$ (Fig. 3).

A well-established correlation between the measured decline $\Delta m_{15}(B)$ of the supernova B-band brightness at maximum and 15 days

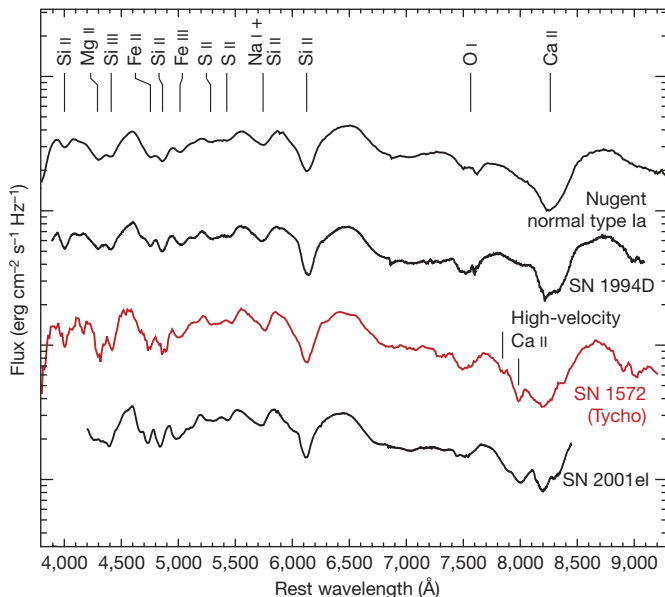


Figure 2 | Spectrum of SN 1572 and comparison spectra of normal type Ia supernovae. Important spectral features are marked. The spectra are plotted logarithmically in flux units and shifted for clarity. The spectrum was obtained with two grisms in the blue ($\lambda < 5000$ Å) and red ($\lambda > 5000$ Å). Total integration time was 4 h: 2.5 h for the red channel and 1.5 h for the blue channel. The spectrum was extracted from a 2.8×2.0 -arcsec 2 aperture (position angle 81°) positioned at the echo brightness peak and then binned to a resolution of 11.2 Å per pixel and smoothed by taking a moving average over five pixels. Flux calibration was performed against the standard star G191B2B, which was observed at comparable airmass. The uncertainty in the flux calibration is 15%. Atmospheric A-band and B-band absorptions were removed using the stellar spectrum of a K star observed in the same slit as the echo. The spectrum was then corrected for the colour dependence of the scattering process for a scattering angle of $\theta = 84^\circ$ and de-reddened for a foreground extinction of $A_V = 4.2$ mag. The scattering angle of $\theta = 84^\circ$ results from the light-echo geometry: because all echo emission at a given epoch is located on an ellipsoidal sphere with the Earth and SN 1572 at its foci, the echo geometry can be accurately determined. For a distance range to the Tycho remnant of 2.3–2.8 kpc, the distance and scattering angle of the echo knot are $d = 460 \pm 45$ light yr and $\theta = 90^\circ \pm 5^\circ$, respectively. For a larger distance of 3.8 kpc, the scattering angle is smaller, $\theta = 67^\circ$, leading to a slightly redder corrected spectrum. However, a slight increase of the adopted foreground extinction by $\Delta A_V = 0.08$ mag compensates for this effect. The comparison spectra have been obtained from the time average of light-curve-weighted spectra at days $-5, -4, -2, +2, +4, +10, +11, +24, +50, +76$ for SN 1994D and days $-9, -4, +1, +9, +18, +40$ for SN 2001el (refs 26, 27).

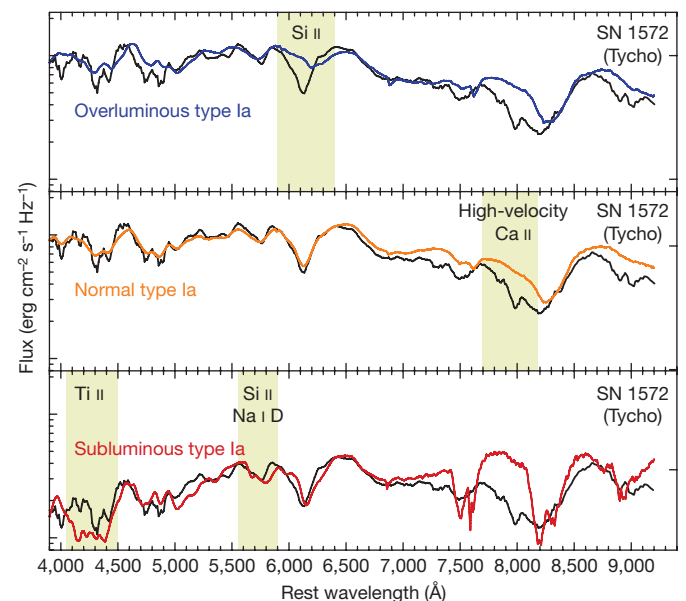


Figure 3 | Comparison of SN 1572 with type Ia supernovae of different luminosities. Spectral templates of subluminous, normal and overluminous type Ia supernovae are shown in comparison with the spectrum of SN 1572. The spectrum of SN 1572 has been corrected for a scattering angle of $\theta = 84^\circ$ and a foreground extinction of $A_V = 4.2$ mag as described for Fig. 2. The comparison spectra have been derived as the time averages of spectral series¹⁹ over days 0–90 after explosion and scaled to the spectrum of SN 1572. Specific features typical of the three subtypes are indicated. For the comparison with the intrinsically redder subluminous template, the spectrum of SN 1572 was de-reddened for a foreground extinction of $A_V = 3.9$ mag.

later has been successfully used to calibrate the absolute brightness of type Ia supernovae²². Applying the range of $1.0 < \Delta m_{15}(B) < 1.3$ obtained from the four type Ia comparison supernovae yields an absolute magnitude of $M_V = -19.0 \pm 0.3$ mag ($H = 72$ km s⁻¹ Mpc⁻¹). Historical records indicate a maximum apparent brightness of between $m_V = -4.0$ mag and $m_V = -4.5$ mag. Accounting for interstellar foreground extinction ($A_V = 1.86 \pm 0.2$ mag (ref. 7)), this places SN 1572 at a distance of $3.8^{+1.5}_{-1.1}$ kpc. This is larger than the currently most quoted range, 2.3–2.8 kpc. However, it is interesting to note that 3.8 kpc is consistent with the non-detection of the remnant at γ -ray energies²³ and the distance to the reported surviving binary companion of SN 1572 (ref. 24). If this discovery of the companion is confirmed it would provide evidence for a single-degenerate scenario in which matter was accreted from this companion.

An interesting difference between the spectrum of SN 1572 and those of normal type Ia supernovae is a deep absorption feature at 7,980 Å in the vicinity of the Ca II infrared triplet (Fig. 2). Whereas normal type Ia supernovae, including SN 1572, show a photospheric absorption minimum of the Ca II infrared triplet at a velocity of 13,500 km s⁻¹, the additional sharp absorption in SN 1572 corresponds to a high velocity component of the Ca II triplet at a velocity of 20,000–24,000 km s⁻¹. Further absorption is detected up to a velocity of 30,000 km s⁻¹. High-velocity components have been shown to be ubiquitous in early spectra of type Ia supernovae²⁵, but they are often mixed with the photospheric absorption. To our knowledge, a feature as strong as that observed in the SN 1572 spectrum has only occasionally been detected in type Ia spectra, for example in SN 2001el. Its spectrum showed a high-velocity Ca II feature at a velocity of 20,000–26,000 km s⁻¹ that was kinematically distinct from the photospheric Ca II absorption^{26,27}. For SN 2001el, spectropolarimetric observations^{26,28} have demonstrated that the high-velocity Ca II feature is the result of an aspherical explosion. Thus, the similarity between the Ca II high-velocity features of SN 1572 and SN 2001el suggests that the asphericity of SN 1572 is also similar to that of SN 2001el. The asphericity could be either due to accretion from a companion or an effect of the explosion.

An exciting opportunity would be to use other SN 1572 light-echo spectra, in different spatial directions, to construct a three-dimensional spectroscopic view of the explosion. This would enable us to determine to what extent spectroscopic diversity can be readily explained by pure geometry effects, using observations of a single supernova. Such observations would further constrain the asphericity suggested by our spectrum. Were the high-velocity Ca II feature caused by a single clump²⁸, no such feature would be expected to be observed in different space directions. Alternatively, the high-velocity Ca II feature could be caused by the interaction with the circumstellar disk of the binary progenitor, in which case the feature might also be observed in other directions²⁹.

An intriguing potential method of determining an independent geometrical distance to the SN 1572 remnant is to search for echoes with the highest degree of polarization and measure their angular distance from the remnant, as has recently been performed for other echoes around a Galactic variable star³⁰. Because the linear polarization of the scattered light is greatest for a scattering angle of 90°, such echoes must be located at a linear distance of ct from the SN 1572 remnant, where c corresponds to the speed of light and t to the time since the original explosion.

Received 21 October; accepted 28 October 2008.

1. Nomoto, K., Thielemann, F.-K. & Yokoi, K. Accreting white dwarf models of Type I supernovae. III - Carbon deflagration supernovae. *Astrophys. J.* **286**, 644–658 (1984).

2. Riess, A. G. *et al.* Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.* **116**, 1009–1038 (1998).
3. Perlmutter, S. *et al.* Measurements of Omega and Lambda from 42 high-redshift supernovae. *Astrophys. J.* **517**, 565–586 (1999).
4. Hillebrandt, W. & Niemeyer, J. C. Type Ia supernova explosion models. *Annu. Rev. Astron. Astrophys.* **38**, 191–230 (2000).
5. Baade, W. B. Cassiopeiae as a supernova of type I. *Astrophys. J.* **102**, 309–317 (1945).
6. Ruiz-Lapuente, P. *et al.* The binary progenitor of Tycho Brahe's 1572 supernova. *Nature* **431**, 1069–1072 (2004).
7. Ruiz-Lapuente, P. Tycho Brahe's supernova: light from centuries past. *Astrophys. J.* **612**, 357–363 (2004).
8. van den Bergh, S. Was Tycho's supernova a subluminal supernova of type Ia? *Astrophys. J.* **413**, 67–69 (1993).
9. Schaefer, B. The peak brightnesses of historical supernovae and the Hubble constant. *Astrophys. J.* **459**, 438–454 (1996).
10. Branch, D. *et al.* Comparative direct analysis of type Ia supernova spectra. II. Maximum light. *Publ. Astron. Soc. Pacif.* **118**, 560–571 (2006).
11. Brahe, T. in *Opera Omnia* Vol. 2 (ed. Dreyer, I. L. E.) 307 (Swets & Zeitlinger, 1972).
12. Doggett, J. B. & Branch, D. A comparative study of supernova light curves. *Astron. J.* **90**, 2303–2311 (1985).
13. Decourchelle, A. *et al.* XMM-Newton observations of the Tycho supernova remnant. *Astron. Astrophys.* **365**, L218–L224 (2001).
14. Krause, O. *et al.* Infrared echoes near the supernova remnant Cas A. *Science* **308**, 1604–1606 (2005).
15. Rest, A. *et al.* Scattered-light echoes from the historical galactic supernovae Cassiopeia A and Tycho (SN 1572). *Astrophys. J.* **681**, L81–L84 (2008).
16. Krause, O. *et al.* The Cassiopeia A supernova was of type IIb. *Science* **320**, 1195–1197 (2008).
17. Rest, A. *et al.* Spectral identification of an ancient supernova using light echoes in the Large Magellanic Cloud. *Astrophys. J.* **680**, 1137–1148 (2008).
18. Filippenko, A. Optical spectra of supernovae. *Annu. Rev. Astron. Astrophys.* **35**, 309–355 (1997).
19. Nugent, P., Kim, A. & Perlmutter, S. K-corrections and extinction corrections for type Ia supernovae. *Publ. Astron. Soc. Pacif.* **114**, 803–819 (2002).
20. Filippenko, A. *et al.* The subluminal, spectroscopically peculiar type Ia supernova 1991bg in the elliptical galaxy NGC 4374. *Astron. J.* **104**, 1543–1556 (1992).
21. Filippenko, A. *et al.* The peculiar type Ia SN 1991T - Detonation of a white dwarf? *Astrophys. J.* **384**, L15–L18 (1992).
22. Phillips, M. M. *et al.* The reddening-free decline rate versus luminosity relationship for type Ia supernovae. *Astron. J.* **118**, 1766–1776 (1999).
23. Völk, H. *et al.* Internal dynamics and particle acceleration in Tycho's SNR. *Astron. Astrophys.* **483**, 529–535 (2008).
24. Gonzalez Hernandez, J. *et al.* The chemical abundances of Tycho G in supernova remnant 1572. *Astrophys. J.* (in the press); preprint at (<http://arxiv.org/abs/0809.0601>) (2008).
25. Mazzali, P. A. *et al.* High-velocity features: A ubiquitous property of type Ia supernovae. *Astrophys. J.* **623**, L37–L40 (2005).
26. Wang, L. *et al.* Spectropolarimetry of SN 2001el in NGC 1448: Asphericity of a normal type Ia supernova. *Astrophys. J.* **591**, 1110–1128 (2003).
27. Mattila, S. *et al.* Early and late time VLT spectroscopy of SN 2001el — progenitor constraints for a type Ia supernova. *Astron. Astrophys.* **443**, 649–662 (2003).
28. Kasen, D. *et al.* Analysis of the flux and polarization spectra of the type Ia supernova SN 2001el: Exploring the geometry of the high-velocity ejecta. *Astrophys. J.* **593**, 788–808 (2003).
29. Tanaka, M. *et al.* Three-dimensional models for high-velocity features in type Ia supernovae. *Astrophys. J.* **645**, 470–479 (2006).
30. Sparks, W. B. *et al.* V838 Monocerotis: a geometric distance from Hubble Space Telescope polarimetric imaging of its light echo. *Astron. J.* **135**, 605–617 (2008).

Acknowledgements This work is based on data collected at the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan, and the German–Spanish Astronomical Center, Calar Alto, jointly operated by the Max-Planck-Institut für Astronomie Heidelberg and the Instituto de Astrofísica de Andalucía. We thank U. Thiele and the Calar Alto observers for their support. M.T. is supported by the Japan Society for the Promotion of Science Research Fellowship for Young Scientists. This research has been supported in part by the World Premier International Research Center Initiative, MEXT, Japan.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.K. (krause@mpia.de) or K.N. (nomoto@astron.s.u-tokyo.ac.jp).

LETTERS

Atmospheric structure and dynamics as the cause of ultraviolet markings in the clouds of Venus

Dmitry V. Titov^{1,4}, Fredric W. Taylor², Håkan Svedhem³, Nikolay I. Ignatiev^{1,4}, Wojciech J. Markiewicz¹, Giuseppe Piccioni⁵ & Pierre Drossart⁶

When seen in ultraviolet light, Venus has contrast features that arise from the non-uniform distribution of unknown absorbers within the sulphuric acid clouds^{1–3} and seem to trace dynamical activity in the middle atmosphere⁴. It has long been unclear whether the global pattern arises from differences in cloud top altitude (which was earlier³ estimated to be 66–72 km), compositional variations or temperature contrasts. Here we report multi-wavelength imaging that reveals that the dark low latitudes are dominated by convective mixing which brings the ultraviolet absorbers up from depth. The bright and uniform mid-latitude clouds reside in the ‘cold collar’, an annulus of cold air characterized by ~30 K lower temperatures with a positive lapse rate, which suppresses vertical mixing and cuts off the supply of ultraviolet absorbers from below. In low and middle latitudes, the visible cloud top is located at a remarkably constant altitude of 72 ± 1 km in both the ultraviolet dark and bright regions, indicating that the brightness variations result from compositional differences caused by the colder environment rather than by elevation changes. The cloud top descends to ~64 km in the eye of the hemispheric vortex, which appears as a depression in the upper cloud deck. The ultraviolet dark circular streaks enclose the vortex eye and are dynamically connected to it.

On 11 April 2006, the Venus Express spacecraft was inserted into an elliptical polar orbit with a 24-h period around Venus^{5,6}. A powerful suite of remote sensing instruments began investigating the morphology, dynamics and conditions at the cloud tops (~70 km), with a view to understanding the meteorology on Venus and comparing it with the Earth⁷. The Venus Monitoring Camera (VMC)⁸ takes wide-angle images in four narrow spectral bands, one of which is centred at 0.365 μm , the characteristic wavelength of the unknown ultraviolet absorber. The images show remarkable variability in the brightness and morphology of the cloud top (Fig. 1). Mottled and spotty clouds at low latitudes indicate vigorous convective activity near the sub-solar point, where most of the solar energy is deposited within the clouds⁹. At about 40° S, they give way to zonally oriented streaks, suggesting that the horizontal non-turbulent flow dominates over convection. Bright and almost featureless clouds prevail in the middle and high latitudes. Poleward of 60° S, global dark circular or spiral jet-like features 200–300 km wide are clearly visible. Despite attempts to explain the observed ultraviolet contrasts using radiative transfer models^{1,2}, the relationship of the ultraviolet features to the physical conditions and dynamical processes at the cloud tops has remained unclear. Even the altitude of formation of the markings was not well determined^{3,10}. Here we constrain the temperature conditions, dynamics and altitude of the cloud tops by using infrared (1–5 μm) observations by the mapping spectrometer VIRTIS¹¹

onboard Venus Express, obtained simultaneously with VMC ultraviolet images, in order to shed light on the origin of the global ultraviolet pattern.

Thermal emission spectra of Venus convey information about the temperature structure above the cloud tops. The shape of the spectra in the 4.3- μm absorption band of CO₂ can be diagnosed in terms of the mesospheric temperature structure (Fig. 2a). Low southern latitudes (<40° S) are characterized by air temperatures that monotonically decrease with altitude (negative lapse rate; Fig. 2b). Middle and high latitudes (>50° S) contain the region with inverted temperature profiles (positive lapse rate) and the coldest conditions at the cloud tops. In earlier observations^{12,13}, researchers discovered a similar structure in the northern hemisphere and named it the ‘cold collar’. Negative lapse rates re-appear in the eye of the hemispheric vortex located poleward of ~75° S. Comparison of the temperature inversion map derived from the VIRTIS observations on the night side with a simultaneously captured VMC ultraviolet image of the

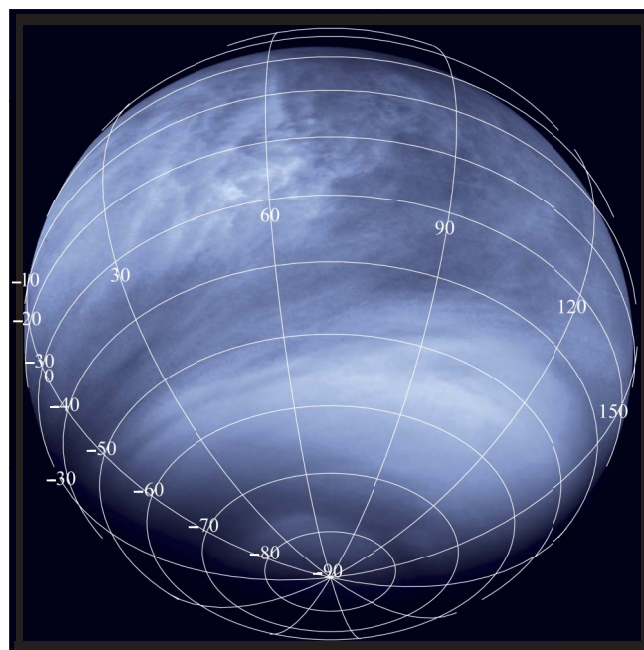


Figure 1 | False-colour image of Venus taken by VMC from a distance of ~30,000 km in the ultraviolet filter. The contrast markings are produced by inhomogeneous spatial and vertical distribution of the unknown ultraviolet absorber mixed within the sulphuric acid cloud.

¹Max Planck Institute for Solar System Research (MPS), Max-Planck-Strasse 2, 37191 Katlenburg-Lindau, Germany. ²University of Oxford, Sub-Department of Atmospheric, Oceanic and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK. ³ESA/ESTEC, PB 299, 2200AG Noordwijk, The Netherlands. ⁴Space Research Institute (IKI), 84/32 Profsoyuznaya Str., 117997 Moscow, Russia. ⁵Istituto di Astrofisica Spaziale e Fisica Cosmica (INAF-IASF), via del Fosso del Cavaliere 100, 00133 Rome, Italy. ⁶LESIA, Observatoire de Paris, 5 place Jules Janssen, 92195 Meudon, France.

day side shows that the location of bright clouds in the ultraviolet image poleward of $\sim 50^\circ$ S correlates with the cold collar (Fig. 2b).

Previous observations placed the ultraviolet markings at 66–72 km with a tentative trend for the cloud top to descend towards the pole^{3,10,13,14}. Multispectral imaging by Venus Express provides an excellent opportunity to map the altitude of the cloud top accurately and to correlate it with the ultraviolet features. The cloud top is located at a remarkably constant altitude of 72 ± 1 km in low and

mid-latitudes (Fig. 3), in good agreement with the earlier polarization studies¹⁰. The cloud top gradually descends from $\sim 60^\circ$ S towards the pole and reaches a minimum of ~ 64 km in the vortex eye. Thus, the altitude of the cloud top varies by about two atmospheric scale heights over the planet. Surprisingly, the sharp outer boundary of the mid-latitude ultraviolet-bright band is not evident in the map of cloud top altitude, implying that both the ultraviolet-dark low latitude and the bright mid-latitude clouds are located at the same altitude level. This suggests that a change in temperature conditions rather than elevation is responsible for the observed global ultraviolet contrasts.

Ultraviolet markings have been routinely used as tracers of the cloud top winds^{4,15}. We find that the observed global distribution of ultraviolet brightness and cloud morphology is also closely related to the dynamical regimes of the lower mesosphere. In particular, the outer edge of the bright mid-latitude band located at $\sim 50^\circ$ S (Fig. 1) marks a transition from low latitudes, where zonal wind is almost constant with latitude, to mid-latitudes where zonal wind quickly fades out towards the pole^{4,9,15,16}. There are tentative indications from the studies of atmospheric chemical tracers¹⁷, primarily carbon monoxide, that this transition also marks the poleward extent of the Hadley cell in the meridional circulation. In addition, tracking of cloud features at several altitudes within the cloud¹⁵ shows that vertical wind shear vanishes poleward of 50° S. The wind field derived from remote temperature sounding using the cyclostrophic approximation shows a mid-latitude jet at 65–70 km and 50° – 55° latitude in both hemispheres¹⁸. Global streaky features in the VMC ultraviolet images apparently indicate the presence of such strong zonal flow around the bright mid-latitude band (Fig. 1).

Another indication of the relation between ultraviolet features and atmospheric dynamics results from the comparison of simultaneous VMC ultraviolet⁹ and VIRTIS thermal-infrared¹⁹ images of the southern polar region (Fig. 4). The vortex eye resides roughly within

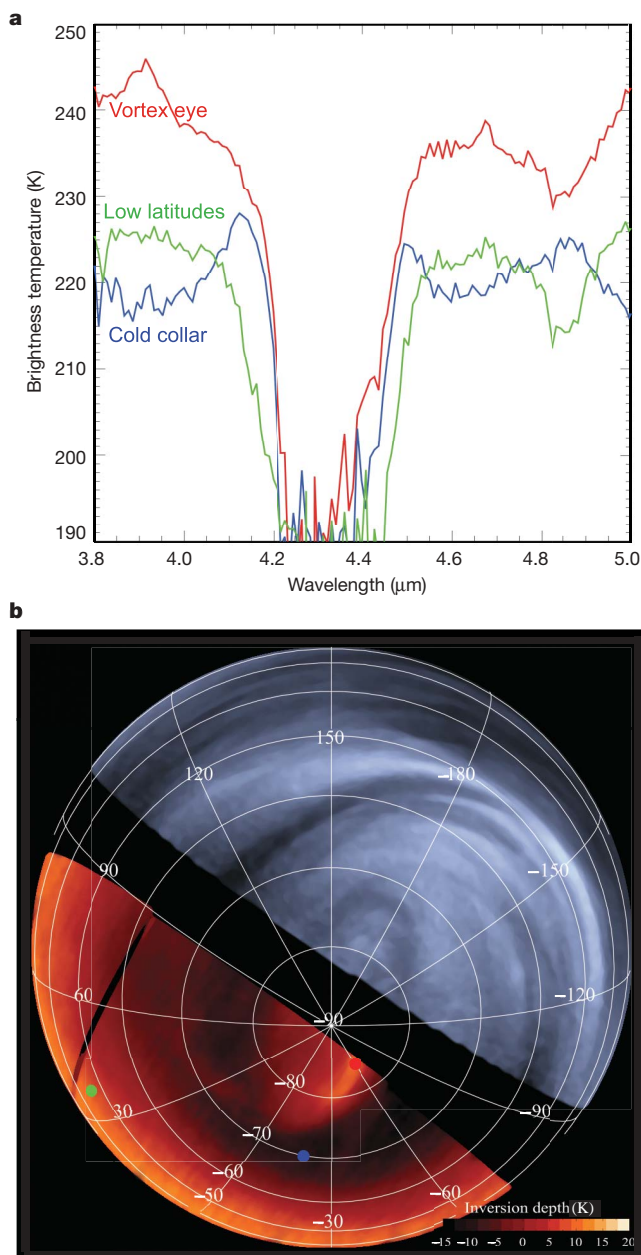


Figure 2 | Thermal infrared spectroscopy of the Venus southern hemisphere by VIRTIS. **a**, Examples of VIRTIS thermal emission spectra in the $4.3\text{-}\mu\text{m}$ CO_2 band for the low latitudes (green), the cold collar (blue) and the eye of the polar vortex (red). Green and red spectra show a monotonic decrease of brightness temperature towards the band centre, which implies a negative lapse rate in the mesosphere. The brightness peaks at $\sim 4.15\text{ }\mu\text{m}$ and $\sim 4.5\text{ }\mu\text{m}$ in the blue spectrum result from an inverted profile of air temperature (positive lapse rate). The peak amplitude—that is, the brightness temperature difference ($\text{BT}_{4.15\text{ }\mu\text{m}} - \text{BT}_{4.0\text{ }\mu\text{m}}$)—is a measure of inversion. **b**, Lower left: a map of the temperature inversion derived from the VIRTIS spectral imaging on the night side. Coloured dots mark the locations corresponding to the spectra in Fig. 2a. Upper right: simultaneously captured VMC ultraviolet image of the day side.

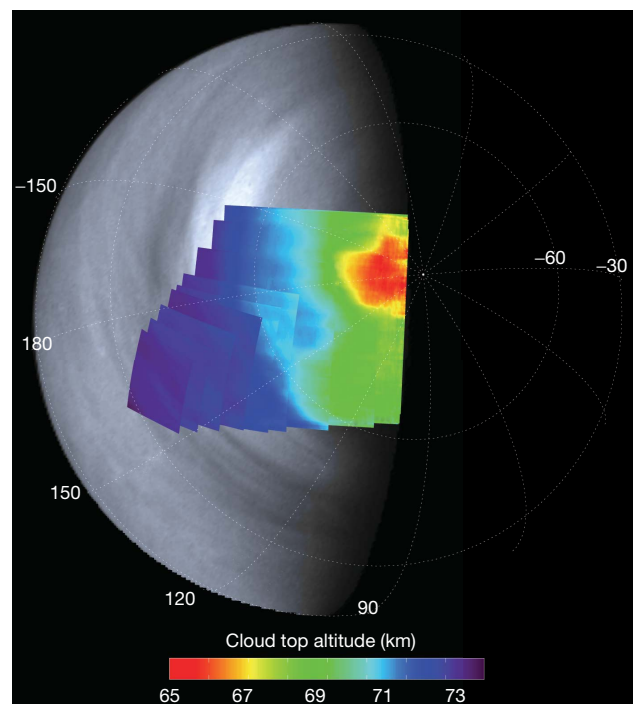


Figure 3 | Altimetry of the cloud tops. The colour mosaic shows a map of the cloud top altitude, derived from VIRTIS spectral imaging in the $1.6\text{-}\mu\text{m}$ CO_2 band, in which relative depth is proportional to the cloud top pressure. The map is plotted on top of a simultaneously captured VMC ultraviolet image. Here we define the cloud top as the level where optical depth $\tau = 1$ at $1.6\text{ }\mu\text{m}$. Radiative transfer calculations show that for the traditional model of the Venus clouds^{1,10} this altitude coincides within a few hundred metres with the level at which $\tau = 1$ at $0.365\text{ }\mu\text{m}$.

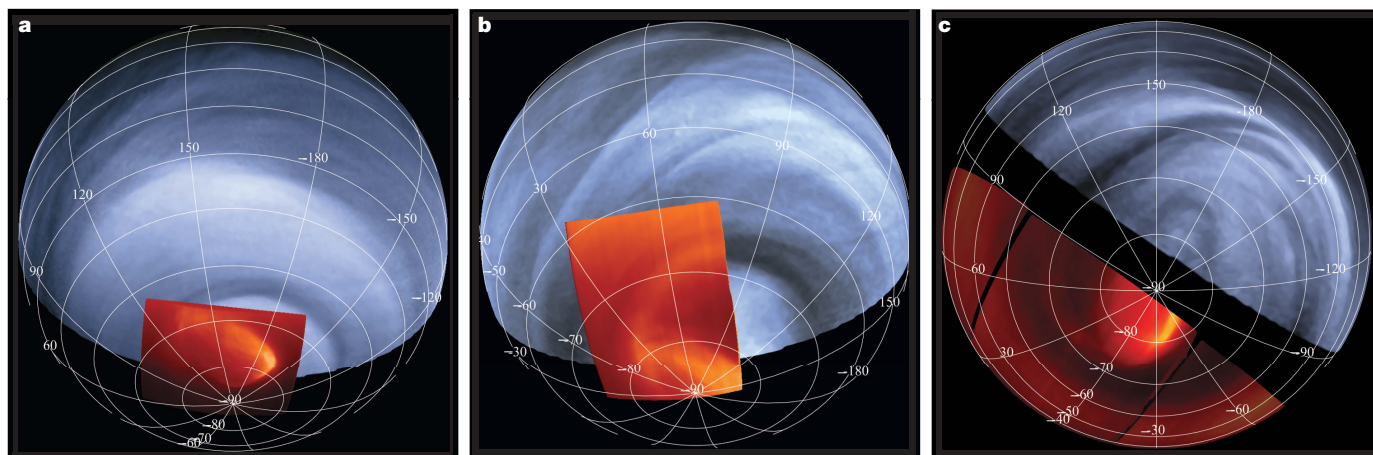


Figure 4 | Composite false colour views of the southern hemisphere. VIRTIS thermal infrared (5- μm) maps are shown as red inserts on top of simultaneously captured VMC ultraviolet (0.365- μm) images. Brightness in the thermal infrared images tracks the temperature of the cloud top. The oval feature in the inserts is the eye of the hemispheric vortex, a dynamical

structure $\sim 2,000$ km in size which is about 30 K warmer than its surroundings. The vortex eye is displaced from the south pole by about 1,000 km, has an irregular and strongly variable shape, and rotates around the pole in about 2.5 days. The atmosphere rotates anticlockwise in the figure.

a 70°S latitude circle, surrounded by an ultraviolet-dark ring (Fig. 4a, b). The spiral arms of the vortex are often connected to this feature. In some cases, the vortex eye or its spiral arms have counterparts in the ultraviolet image¹⁹ (Fig. 4c). All of this suggests that ultraviolet and thermal infrared features observed at the pole are both manifestations of the same dynamical phenomenon in the polar mesosphere.

A consistent picture of the global ultraviolet pattern of the cloud tops and its relation to the physical conditions and dynamics of the lower mesosphere is emerging from the multispectral imaging by VIRTIS and VMC onboard Venus Express. Three latitude zones can be distinguished on the planet (Fig. 5): ultraviolet-dark low latitudes ($<50^\circ\text{S}$), a ultraviolet-bright mid-latitude band coinciding with the cold collar ($50^\circ\text{--}70^\circ\text{S}$), and a polar 'cap' ($>70^\circ\text{S}$) with embedded vortex eye. They differ considerably in cloud morphology, ultraviolet appearance, temperature structure and dynamics.

Strong convection induced by deposition of solar energy inside the main cloud^{20,21} dominates in low latitudes. We see its traces clearly in the cloud-top morphology⁹. Convective mixing brings ultraviolet absorbers from depth, making low latitudes appear relatively dark at these wavelengths. This explanation qualitatively agrees with the dynamical scheme of ultraviolet contrast formation proposed by Esposito and Travis¹⁰.

The situation changes markedly at $\sim 50^\circ\text{S}$ when we enter the cold collar, an annulus of cold air located right at the cloud tops. Thermal emission spectroscopy by VIRTIS¹⁹ (Fig. 2) and radio-occultation sounding by VeRa/Venus Express²² suggest the presence of strong temperature inversions here. This in turn implies high convective stability and suppressed vertical mixing. The dynamical state in the

cold collar also differs remarkably from that in low latitudes. Streaky clouds point to non-turbulent horizontal flow that dominates over convection (Fig. 1). Both zonal wind velocity and its vertical gradient quickly decrease with latitude^{4,9,15,16}, preventing development of shear instabilities. All this suppresses the supply of absorbers from depth and explains the bright appearance and the paucity of ultraviolet features in middle latitudes. In addition, the cold temperatures in the collar region create favourable conditions for formation of bright sulphuric acid haze. This dense conservatively scattering aerosol masks the ultraviolet-absorbing layer hidden deeper inside the cloud.

The vortex eye resides inside the $\sim 70^\circ$ latitude circle marked by the ultraviolet-dark polar ring (Fig. 4). The air temperature here gradually increases with depth, in contrast with the surroundings, which show well-developed inversions. We discover that the vortex eye is apparently a deep depression in the cloud top, which is located at about 64 km here, that is, ~ 8 km or two atmospheric scale heights lower than anywhere else on the planet (Fig. 3). This deep location of the clouds, maintained by dynamics, together with monotonically increasing air temperature, explains the high brightness temperature observed in the vortex eye. The global morphology of the cloud tops in the southern hemisphere studied in detail by Venus Express is similar to that discovered by Pioneer Venus¹² and Venera-15¹³ in the north, thus suggesting global hemispherical symmetry¹⁹.

The vortex eye is located inside the $\sim 70^\circ$ latitude circle marked by the ultraviolet-dark polar ring (Fig. 4). The air temperature here gradually increases with depth, in contrast with the surroundings, which show well-developed inversions. We discover that the vortex eye is apparently a deep depression in the cloud top, which is located at about 64 km here, that is, ~ 8 km or two atmospheric scale heights lower than anywhere else on the planet (Fig. 3). This deep location of the clouds, maintained by dynamics, together with monotonically increasing air temperature, explains the high brightness temperature observed in the vortex eye. The global morphology of the cloud tops in the southern hemisphere studied in detail by Venus Express is similar to that discovered by Pioneer Venus¹² and Venera-15¹³ in the north, thus suggesting global hemispherical symmetry¹⁹.

Received 4 April; accepted 16 September 2008.

- Pollack, J. B. *et al.* Distribution and source of the UV absorption in Venus' atmosphere. *J. Geophys. Res.* **85**, 8141–8150 (1980).
- Esposito, L. W. Ultraviolet contrasts and the absorbers near the Venus cloud tops. *J. Geophys. Res.* **85**, 8151–8157 (1980).
- Esposito, L. W. *et al.* in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 484–564 (Univ. Arizona Press, 1983).
- Rossow, W. B. *et al.* Cloud morphology and motions from Pioneer Venus images. *J. Geophys. Res.* **85**, 8107–8128 (1980).
- Svedhem, H. *et al.* Venus Express—the first European mission to Venus. *Planet. Space Sci.* **55**, 1636–1652 (2007).
- Titov, D. V. *et al.* Venus Express science planning. *Planet. Space Sci.* **54**, 1279–1297 (2006).
- Taylor, F. W., Svedhem, H. & Titov, D. in *Exploring Venus as a Terrestrial Planet* (eds Esposito, L. W., Stofan, E. R. & Cravens, T. E.) 157–170 (Geophysical Monograph 176, American Geophysical Union, 2007).
- Markiewicz, W. J. *et al.* Venus Monitoring Camera for Venus Express. *Planet. Space Sci.* **55**, 1701–1711 (2007).
- Markiewicz, W. J. *et al.* Morphology and dynamics of the upper cloud layer of Venus. *Nature* **450**, 633–636 (2007).
- Esposito, L. W. & Travis, L. D. Polarization studies of the Venus UV contrasts: cloud height and haze variability. *Icarus* **51**, 374–390 (1982).
- Drossart, P. *et al.* Scientific goals for the observations of Venus by VIRTIS on ESA/Venus Express mission. *Planet. Space Sci.* **55**, 1653–1672 (2007).

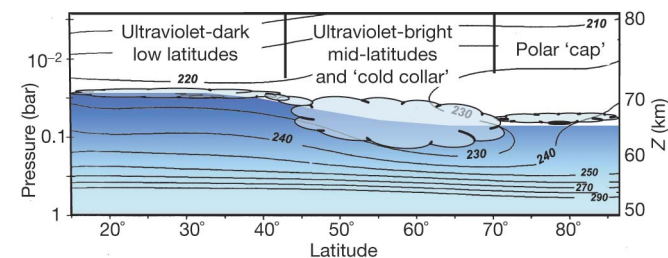


Figure 5 | Sketch of the global morphology of the cloud top. Isolines show the mesospheric temperature field derived from Venera-15 spectrometry in the northern hemisphere¹⁸. Conservatively scattering ultraviolet-bright cloud is shown in light blue, and the ultraviolet-absorbing layer is in dark blue.

12. Taylor, F. W. *et al.* Structure and meteorology of the middle atmosphere of Venus: infrared remote sensing from the Pioneer orbiter. *J. Geophys. Res.* **85**, 7963–8006 (1980).
 13. Zasova, L. V., Ignatiev, N., Khatuntsev, I. & Linkin, V. Structure of the Venus atmosphere. *Planet. Space Sci.* **55**, 1712–1728 (2007).
 14. Zasova, L. V., Moroz, V. I., Esposito, L. W. & Na, C. Y. SO₂ in the middle atmosphere of Venus: IR measurements from Venera-15 and comparison to UV data. *Icarus* **105**, 92–109 (1993).
 15. Sánchez-Lavega, A. *et al.* Variable winds on Venus mapped in three dimensions. *Geophys. Res. Lett.* **35**, L13204, doi:10.1029/2008GL033817 (2008).
 16. Limaye, S. S. Venus atmospheric circulation: Known and unknown. *J. Geophys. Res.* **112**, E04S09, doi:10.1029/2006JE002814 (2007).
 17. Tsang, C. C. C. *et al.* Tropospheric carbon monoxide concentrations and variability on Venus from Venus Express/VIRTIS-M observations. *J. Geophys. Res.* doi:10.1029/2008JE003089 (2008).
 18. Lellouch, E. *et al.* in *Venus II* (eds Bougher, S. W., Hunten, D. M. & Phillips, R. J.) 295–324 (Univ. Arizona Press, 1997).
 19. Piccioni, G. *et al.* South-polar features on Venus similar to those near the north pole. *Nature* **450**, 637–640 (2007).
 20. Crisp, D. & Titov, D. V. in *Venus II* (eds Bougher, S. W., Hunten, D. M. & Phillips, R. J.) 353–384 (Univ. Arizona Press, 1997).
 21. Titov, D. V. *et al.* in *Exploring Venus as a Terrestrial Planet* (eds Esposito, L. W., Stofan, E. R. & Cravens, T. E.) 121–138 (Geophysical Monograph 176, American Geophysical Union, 2007).
 22. Pätzold, M. *et al.* The structure of Venus' middle atmosphere and ionosphere. *Nature* **450**, 657–660 (2007).
- Acknowledgements** We thank the European Space Agency (ESA) for its efforts and dedication in implementation of the Venus Express mission and the teams at ESOC and ESAC for operational support. We thank the UK Science and Technology Facilities Council, Italian (ASI) and French (CNES) space agencies for support. N.I.I. acknowledges support from the Russian Foundation for Basic Research.
- Author Contributions** D.V.T. led the work and studied correlations of the ultraviolet and thermal infrared features in VMC and VIRTIS images; F.W.T. worked on meteorological aspects; H.S. is the Venus Express Project Scientist; N.I.I. derived the maps of cloud top altitude from the VIRTIS spectra; W.J.M., G.P. and P.D. are the Principal Investigators of VMC and VIRTIS experiments and led the observations. All authors discussed the results and commented on the manuscript.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.V.T. (Titov@mps.mpg.de).

LETTERS

Electric field control of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface ground state

A. D. Caviglia¹, S. Gariglio¹, N. Reyren¹, D. Jaccard¹, T. Schneider², M. Gabay³, S. Thiel⁴, G. Hammerl⁴, J. Mannhart⁴ & J.-M. Triscone¹

Interfaces between complex oxides are emerging as one of the most interesting systems in condensed matter physics¹. In this special setting, in which translational symmetry is artificially broken, a variety of new and unusual electronic phases can be promoted². Theoretical studies predict complex phase diagrams and suggest the key role of the charge carrier density in determining the systems' ground states. A particularly fascinating system is the conducting interface between the band insulators LaAlO_3 and SrTiO_3 (ref. 3). Recently two possible ground states have been experimentally identified: a magnetic state⁴ and a two-dimensional superconducting condensate⁵. Here we use the electric field effect to explore the phase diagram of the system. The electrostatic tuning of the carrier density allows an on/off switching of superconductivity and drives a quantum phase transition^{6–8} between a two-dimensional superconducting state and an insulating state. Analyses of the magnetotransport properties in the insulating state are consistent with weak localization and do not provide evidence for magnetism. The electric field control of superconductivity demonstrated here opens the way to the development of new mesoscopic superconducting circuits.

Since the discovery of conductivity at the $\text{LaAlO}_3/\text{SrTiO}_3$ interface, one of the main challenges has been the identification of the source of charge carriers. Despite much research effort^{9–12} there is not yet a general consensus on the nature of the doping mechanism. Depending on the growth conditions, the doping can be related to the polar nature of the LaAlO_3 atomic planes¹³ (the polar catastrophe model or 'intrinsic doping'), to the creation of oxygen defects during the growth of the samples^{9–11} or to interdiffusion phenomena¹² ('extrinsic doping'). It is generally accepted that the conduction observed in samples grown at low oxygen pressure (less than about 10^{-6} mbar) is dominated by oxygen defects and extends deeply into the substrate, whereas for samples grown at higher pressure (over about 10^{-5} mbar) the conduction is confined to the interface¹⁴. Both superconductivity⁵ and unusual magnetotransport properties, attributed to magnetism⁴, have recently been observed in samples grown in the high-pressure regime. These results suggest that the ground state of the system might be sensitive to small changes in carrier concentration, or to the amount of disorder, or both. To resolve this issue and investigate the phase diagram of the system, electrostatic doping seems to be an ideal technique, because it allows the tuning of the carrier density while preserving the oxygen concentration and disorder landscape^{15,16}. Recent reports indicate that an electric field can effectively modulate the transport properties of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface¹⁷.

$\text{LaAlO}_3/\text{SrTiO}_3$ superconducting interfaces and field-effect devices were prepared as described in the Methods. In a standard field-effect

device, an electric field is applied between a metallic gate and a conducting channel across a dielectric. We used the 0.5-mm-thick SrTiO_3 substrate as the dielectric because it is characterized at low temperatures by a large dielectric constant. The metallic gate is a gold film sputtered opposite to the channel area onto the back of the substrate. In this configuration the electric field modulates the concentration of carriers in the interface conducting channel. A sketch of the field-effect device is shown in Fig. 1.

To quantify the carrier density modulation resulting from the field effect, the electric field dependence of the SrTiO_3 permeability needs to be taken into account¹⁸. We therefore measured the differential capacitance $C(V) = dQ(V)/dV$ of the device as a function of the applied gate voltage V , $Q(V)$ being the induced charge. Charge trapping can occur in SrTiO_3 and may cause the appearance of hysteresis in the $C(V)$ measurements¹⁹. Indeed, we found that the capacitance depended on the voltage sweep history. However, the devices present reversible and reproducible $C(V)$ characteristics when the field is first ramped to the highest positive voltage. Following this experimental procedure, the field-induced modulation of charge density δn_{2D} between gate voltages V_1 and V_2 can be evaluated using the relation

$$\delta n_{2D} = \frac{1}{Se} \int_{V_1}^{V_2} C(V) dV \quad (1)$$

where S is the area of the gate electrode and e is the elementary charge. The $C(V)$ characteristic of a device made of a layer of LaAlO_3 nine unit

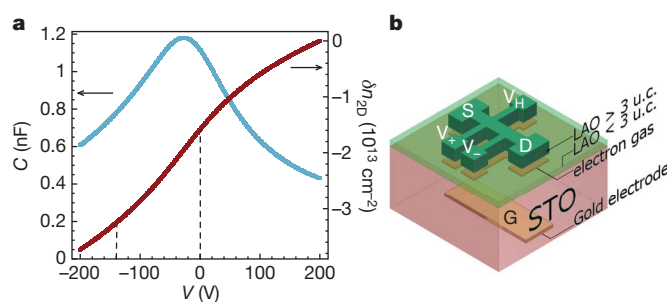


Figure 1 | Dielectric characterization of the field-effect device. **a**, Dielectric tunability of the differential capacitance (C versus V) measured on a device at 1.5 K (blue). The capacitance $C(V)$ is measured with an a.c. technique applying a signal amplitude of 1 V. Change in the 2D carrier concentration as a function of gate voltage is calculated using equation (1) with $V_1 = 200$ V and $V_2 = V$ (red). The dashed lines indicate the region where the quantum critical behaviour is observed. Note that in this region $\delta n_{2D} \propto \delta V$.

b, Schematic view of a field-effect device, showing the source (S), drain (D), longitudinal voltage (V_+ and V_-), Hall voltage (V_H) and gate voltage (G) contacts. STO, SrTiO_3 ; LAO, LaAlO_3 .

¹Département de Physique de la Matière Condensée, University of Geneva, 24 Quai Ernest-Ansermet, 1211 Genève 4, Switzerland. ²Physikinstitut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. ³Laboratoire de Physique des Solides, Bat 510, Université Paris-Sud 11, Centre d'Orsay, 91405 Orsay Cedex, France. ⁴Experimental Physics VI, Center for Electronic Correlations and Magnetism, Institute of Physics, University of Augsburg, D-86135 Augsburg, Germany.

cells thick, and the corresponding modulation of carrier density, are presented in Fig. 1a. The carrier concentration of the as-grown sample has been measured using the Hall effect ($n_{2D} \approx 4.5 \times 10^{13} \text{ cm}^{-2}$ at 100 K). The maximum modulation of the carrier density that was achieved is remarkably close to the total number of free carriers present in the system, indicating that the electric field effect is an excellent tool to probe its phase diagram. On the same sample we measured the temperature dependence of the sheet resistance R_{sheet} for various gate voltages V .

Figure 2a shows the sheet resistance as a function of temperature for applied gate voltages between -300 V and 320 V ; Fig. 2b displays the same data on a linear sheet-resistance scale. This behaviour has been observed in several samples. A variation of the gate voltage induces a large modulation of the normal-state resistance, which changes by two orders of magnitude, and a remarkable tuning of the superconducting critical temperature. For large negative voltages, corresponding to the smallest accessible electron densities, the sheet resistance increases as the temperature is decreased, indicating an insulating ground state (conductance $G \rightarrow 0$ as $T \rightarrow 0$). As the electron density is increased the system becomes a superconductor. The transition from superconducting to insulating ground state occurs at a critical sheet resistance of $R_c \approx 4.5 \text{ k}\Omega$ per square, close to the quantum of resistance for bosons with charge $2e$, $R_Q = h/4e^2 \approx 6.45 \text{ k}\Omega$. A further increase in the electron density produces first a rise of the critical temperature to a maximum of $\sim 310 \text{ mK}$. For larger voltages, the critical temperature decreases again, providing evidence for an overdoped regime. These measurements thus reveal the existence of a quantum phase transition between a superconducting and an insulating phase at the $\text{LaAlO}_3/\text{SrTiO}_3$ interface and demonstrate that the ground state of the system depends on its carrier density.

Close to the critical point, the sheet resistance measured at 400 mK shows a remarkable phenomenon. As can be seen in Fig. 3, approaching the quantum critical point from the superconducting region of the phase diagram, a linear dependence of the sheet resistance as a function of the applied voltage is observed. Once the critical point has been crossed, however, a further reduction of carrier concentration produces a much steeper variation in resistance.

To establish the critical temperature versus carrier density phase diagram, a criterion to define the critical temperature for each gate voltage is needed. It has been shown that the superconducting transitions are consistent with the Berezinskii–Kosterlitz–Thouless (BKT)

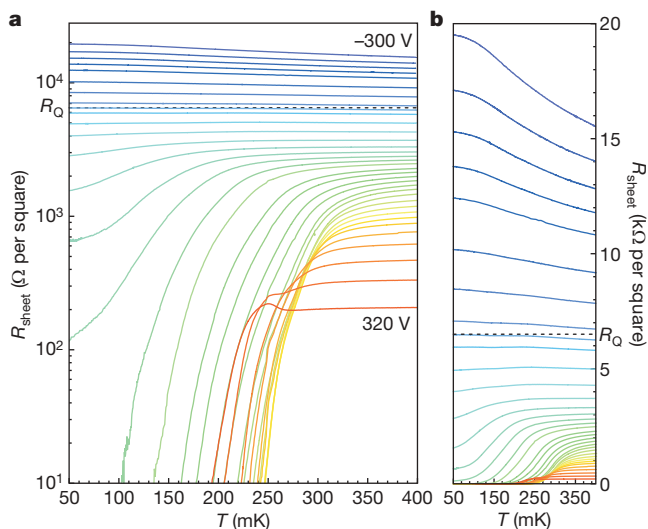


Figure 2 | Field-effect modulation of the transport properties. **a**, Measured sheet resistance as a function of temperature, plotted on a semi-logarithmic scale, for gate voltages varying in 10-V steps between -300 V and -260 V , 20-V steps between -260 V and 320 V , and for -190 V . The dashed line indicates the quantum of resistance R_Q . **b**, The same data plotted on a linear resistance scale.

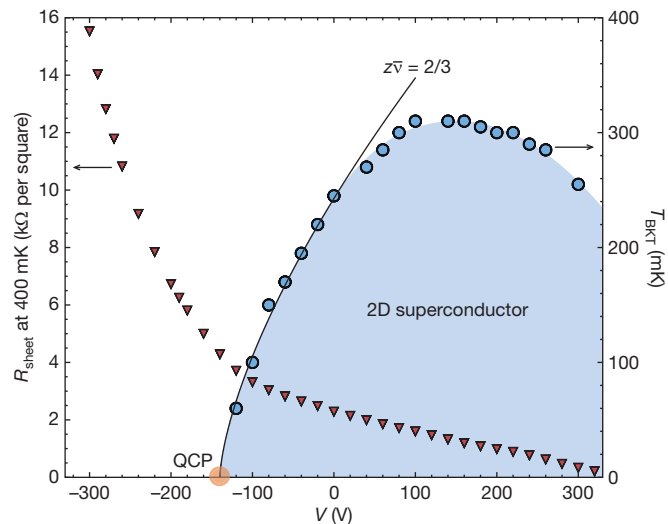


Figure 3 | Electronic phase diagram of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface. Critical temperature T_{BKT} (right axis, blue dots) is plotted against gate voltage, revealing the superconducting region of the phase diagram. The solid line describes the approach to the quantum critical point (QCP) using the scaling relation $T_{\text{BKT}} \propto (V - V_c)^{z\nu}$, with $z\nu = 2/3$. Also plotted is normal-state sheet resistance, measured at 400 mK (left axis, red triangles) as a function of gate voltage.

behaviour expected for a two-dimensional system⁵. According to the BKT model, above the critical temperature T_{BKT} the temperature dependence of the resistance is expected to be

$$R \propto \exp\left(-\frac{b_R}{(T - T_{\text{BKT}})^{1/2}}\right) \quad (2)$$

where b_R is a parameter related to the vortex properties.

This approach allows T_{BKT} to be extracted for each applied gate voltage and the phase diagram to be mapped. The result is shown in Fig. 3. Reducing the carrier concentration from the largest doping level ($V = 320 \text{ V}$), T_{BKT} first increases, reaches a maximum at around 310 mK and then decreases to zero. This critical line ends at $V_c \approx -140 \text{ V}$, where the system undergoes a quantum phase transition (QPT).

To investigate the quantum critical region in detail we consider a continuous QPT separating a superconducting ground state and an insulating ground state. The control parameter of the phase transition is the variation of the carrier concentration $\delta n_{2D} = n_{2D} - n_{2Dc}$, where n_{2Dc} is the sheet carrier density at the critical point. Figure 1a shows a quasi-linear relationship between the applied gate voltage and the variation of carrier concentration $\delta n_{2D} \propto \delta V = V - V_c$ near the critical point. Hence we can use the gate voltage as the tuning parameter for the analysis of the QPT. For a continuous QPT, the quantum critical region is characterized by a spatial and a temporal correlation length that diverge respectively as $\xi \propto (\delta n_{2D})^{-\nu}$ and $\xi_\tau \propto (\delta n_{2D})^{-\nu}$. The quantum dynamic critical exponent is defined through the ratio $z = \nu_\tau/\nu$. According to the scaling theory of quantum critical phenomena^{20–22} the phase transition line $T_{\text{BKT}}(\delta V)$ presented in Fig. 3 is expected to scale, near the quantum critical point, as

$$T_{\text{BKT}} \propto (\delta n_{2D})^{z\nu} \propto (\delta V)^{z\nu} \quad (3)$$

In Fig. 3 we observe that the approach to quantum criticality is well described by $z\nu = 2/3$, exponents that also describe well $T_{\text{BKT}} \propto (\delta n_{2D})^{z\nu}$. We note that the product $z\nu = 2/3$ agrees with the results obtained in previous studies of two-dimensional (2D) quantum superconductor to insulator phase transitions driven by magnetic and electric fields in amorphous bismuth films²³ and $\text{Nb}_{0.15}\text{Si}_{0.85}$ films²⁴. The value $z\nu = 2/3$ is compatible with the 3D-XY model⁶,

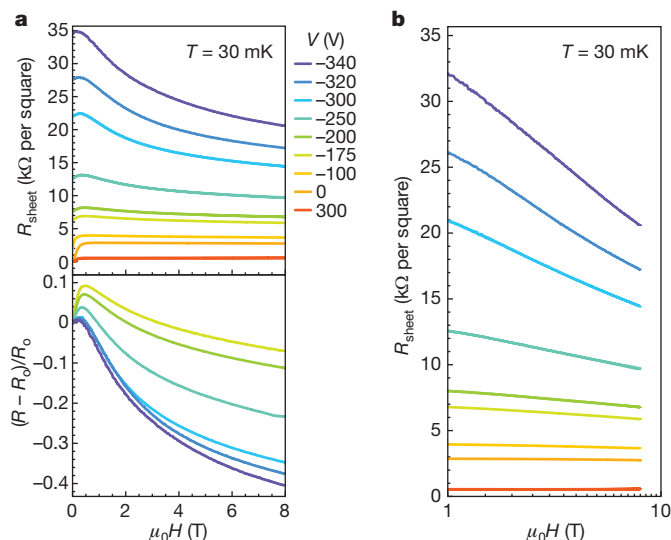


Figure 4 | Field-effect modulation of the magnetotransport properties. **a**, Sheet resistance (upper panel) and magnetoresistance (lower panel), as a function of magnetic field, measured for different gate voltages at 30 mK. Note the large negative magnetoresistance for the lowest carrier densities. **b**, Sheet resistance as a function of magnetic field, measured for different gate voltages at 30 mK and plotted on a logarithmic magnetic field scale.

possibly indicative of a clean (or weakly disordered) 2D system in which quantum fluctuations dominate ((2+1)D-XY). The voltage range over which critical fluctuations can be observed has, however, to be determined. We note that a highly disordered (dirty boson) system²⁵ with localized Cooper pairs in the insulating phase would yield $z\nu = 1$. This value would agree with the critical behaviour observed in $\text{NdBa}_2\text{Cu}_3\text{O}_{7-\delta}$ thin films²⁶. Although the dirty boson interpretation is not incompatible with the data, in most systems dominated by disorder the magnetoresistance is expected to be positive, at variance with our observations (see below).

Of particular relevance for the understanding of the nature of this electronic system is the insulating region of the phase diagram. In the accessible temperature range the variation of the conductance can be explained by weak localization^{27,28}. Figure 4a depicts the dependence of the sheet resistance in a perpendicular magnetic field measured at 30 mK. As expected if weak localization is governing the magnetotransport properties, we observe a large negative magnetoresistance that increases as we move more deeply into the insulating phase, reaching more than -40% at 8 T for the lowest measured carrier concentration. Above ~ 1 T the resistance decreases logarithmically with increasing magnetic field (Fig. 4b). In the non-superconducting region of the phase diagram no hysteresis in magnetoresistance has been detected (see also Methods). Our data are in partial agreement with the results presented in ref. 4, where a negative magnetoresistance was also reported.

The field-effect tuning of the electronic properties of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface achieved here unravels a complex phase diagram controlled by the charge density with a quantum critical point separating a superconducting and an insulating region. Using the electrostatic control of superconductivity, new quantum electronic devices can be envisaged where, for instance, superconductivity can be dynamically defined and controlled using local electric fields.

METHODS SUMMARY

We prepared conducting interfaces by depositing LaAlO_3 films, with thickness more than four unit cells, on TiO_2 -terminated (001) SrTiO_3 single crystals. The films were grown by pulsed laser deposition at $\sim 800^\circ\text{C}$ in $\sim 1 \times 10^{-4}$ mbar of O_2 with a repetition rate of 1 Hz. The fluence of the laser pulses was 0.6 J cm^{-2} . We monitored film growth *in situ* using reflection high-energy electron diffraction which allowed the thickness to be controlled with sub-unit-cell precision²⁹. After

growth, each sample was annealed in 200 mbar of O_2 at about 600°C for one hour and cooled to room temperature in the same oxygen pressure. Samples with different thickness of LaAlO_3 were systematically characterized using atomic force microscopy and X-ray diffraction. The samples were then patterned in a geometry suitable for four-point transport measurements. The patterning technique is based on the discovery that only regions covered by a LaAlO_3 layer with thickness ≥ 4 unit cells are conducting¹⁷. The LaAlO_3 thickness was thus reduced to about two unit cells or less in specific regions of the samples, irradiating them with argon ions while protecting transport channels with a photoresist layer. The superconducting interface was contacted by ultrasonic welding using Al-wires. Transport properties were measured in a dilution cryostat applying a linear current density of $2 \times 10^{-4} \text{ A m}^{-1}$ in a $500\text{-}\mu\text{m}$ -wide path. Magnetic fields were applied perpendicular to the sample and ramped with a sweep rate of 1.6 mT s^{-1} . Higher sweep rates (3.2 mT s^{-1}) revealed, in the superconducting region of the phase diagram, small hystereses (1%) in the magnetoresistance. Such hystereses were not observed in the insulating region. Electric fields were applied using a voltage source. No leakage currents ($\leq 1 \text{ nA}$) could be detected up to the largest applied voltages. The devices present reversible and reproducible $C(V)$ characteristics when the field is first ramped to the highest positive voltage. Note that in our experimental procedure the 0 V state does not correspond to the as-grown state. The maximum applied electric field was chosen to avoid any breakdown of the electrical insulation inside the dilution cryostat.

Received 1 July; accepted 20 October 2008.

- Hwang, H. Y. Atomic control of the electronic structure at complex oxide heterointerfaces. *Mater. Res. Soc. Bull.* **31**, 28–35 (2006).
- Okamoto, S. & Millis, A. J. Electronic reconstruction at an interface between a Mott insulator and a band insulator. *Nature* **428**, 630–633 (2004).
- Ohtomo, A. & Hwang, H. Y. A high-mobility electron gas at the $\text{LaAlO}_3/\text{SrTiO}_3$ heterointerface. *Nature* **427**, 423–426 (2004).
- Brinkman, A. *et al.* Magnetic effects at the interface between non-magnetic oxides. *Nature Mater.* **6**, 493–496 (2007).
- Reyren, N. *et al.* Superconducting interfaces between insulating oxides. *Science* **317**, 1196–1199 (2007).
- Sachdev, S. *Quantum Phase Transitions* (Cambridge Univ. Press, 1999).
- Sondhi, S. L., Girvin, S. M., Carini, J. P. & Shahar, D. Continuous quantum phase transitions. *Rev. Mod. Phys.* **69**, 315–333 (1997).
- v. Löhneysen, H., Rosch, A., Vojta, M. & Wölfle, P. Fermi-liquid instabilities at magnetic quantum phase transitions. *Rev. Mod. Phys.* **79**, 1015–1075 (2007).
- Herranz, G. *et al.* High mobility in $\text{LaAlO}_3/\text{SrTiO}_3$ heterostructures: origin, dimensionality, and perspectives. *Phys. Rev. Lett.* **98**, 216803 (2007).
- Kalabukhov, A. *et al.* Effect of oxygen vacancies in the SrTiO_3 substrate on the electrical properties of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface. *Phys. Rev. B* **75**, 121404 (2007).
- Siemons, W. *et al.* Origin of charge density at LaAlO_3 on SrTiO_3 heterointerfaces: Possibility of intrinsic doping. *Phys. Rev. Lett.* **98**, 196802 (2007).
- Willmott, P. R. *et al.* Structural basis for the conducting interface between LaAlO_3 and SrTiO_3 . *Phys. Rev. Lett.* **99**, 155502 (2007).
- Nakagawa, N., Hwang, H. Y. & Muller, D. A. Why some interfaces cannot be sharp. *Nature Mater.* **5**, 204–209 (2006).
- Basletic, M. *et al.* Mapping the spatial distribution of charge carriers in $\text{LaAlO}_3/\text{SrTiO}_3$ heterostructures. *Nature Mater.* **7**, 621–625 (2008).
- Ahn, C. H., Triscone, J.-M. & Mannhart, J. Electric field effect in correlated oxide systems. *Nature* **424**, 1015–1018 (2003).
- Ahn, C. H. *et al.* Electrostatic modification of novel materials. *Rev. Mod. Phys.* **78**, 1185–1212 (2006).
- Thiel, S., Hammerl, G., Schmehl, A., Schneider, C. W. & Mannhart, J. Tunable quasi-two-dimensional electron gases in oxide heterostructures. *Science* **313**, 1942–1945 (2006).
- Matthey, D., Gariglio, S. & Triscone, J.-M. Field-effect experiments in $\text{NdBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ultrathin films using a SrTiO_3 single-crystal gate insulator. *Appl. Phys. Lett.* **83**, 3758–3760 (2003).
- Christen, H.-M., Mannhart, J., Williams, E. J. & Gerber, C. Dielectric properties of sputtered SrTiO_3 films. *Phys. Rev. B* **49**, 12095–12104 (1994).
- Schneider, T. Universal critical quantum properties of cuprate superconductors. *Acta Phys. Polon. A* **91**, 203–212 (1997).
- Schneider, T. *The Physics of Superconductors* Vol. 2, 111 (Springer, 2004).
- Schneider, T. & Singer, J. M. *Phase Transition Approach to High Temperature Superconductivity* (Imperial College Press, 2000).
- Parendo, K. A. *et al.* Electrostatic tuning of the superconductor–insulator transition in two dimensions. *Phys. Rev. Lett.* **94**, 197004 (2005).
- Aubin, H. *et al.* Magnetic-field-induced quantum superconductor–insulator transition in $\text{Nb}_{0.15}\text{Si}_{0.85}$. *Phys. Rev. B* **73**, 094521 (2006).
- Fisher, M. P. A. & Grinstein, G. Quantum critical phenomena in charged superconductors. *Phys. Rev. Lett.* **60**, 208–211 (1988).
- Matthey, D., Reyren, N., Schneider, T. & Triscone, J.-M. Electric-field-effect modulation of the transition temperature, mobile carrier density, and in-plane penetration depth of $\text{NdBa}_2\text{Cu}_3\text{O}_{7-\delta}$ thin films. *Phys. Rev. Lett.* **98**, 057002 (2007).

27. Bergmann, G. Weak localization in thin films. *Phys. Rep.* **107**, 1–58 (1984).
28. Das, D. & Doniach, S. Weakly localized bosons. *Phys. Rev. B* **57**, 14440–14443 (1998).
29. Rijnders, G., Koster, G., Blank, D. H. A. & Rogalla, H. In situ monitoring during pulsed laser deposition of complex oxides using reflection high energy electron diffraction under high oxygen pressure. *Appl. Phys. Lett.* **70**, 1888–1890 (1997).

Acknowledgements We thank T. Giamarchi, L. Benfatto, T. Kopp and A.-S. Ruetschi for discussions and M. Lopes for technical assistance. We acknowledge

financial support by the Swiss National Science Foundation through the National Centre of Competence in Research ‘Materials with Novel Electronic Properties’ MaNEP and Division II, by the European Union through the project ‘Nanoxide’, by the Deutsche Forschungsgemeinschaft through the SFB484, and by the European Science Foundation through the program ‘Thin Films for Novel Oxide Devices’.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.D.C. (andrea.caviglia@unige.ch).

LETTERS

Large tundra methane burst during onset of freezing

Mikhail Mastepanov¹, Charlotte Sigsgaard², Edward J. Dlugokencky³, Sander Houweling^{4,5}, Lena Ström¹, Mikkel P. Tamstorf⁶ & Torben R. Christensen¹

Terrestrial wetland emissions are the largest single source of the greenhouse gas methane¹. Northern high-latitude wetlands contribute significantly to the overall methane emissions from wetlands, but the relative source distribution between tropical and high-latitude wetlands remains uncertain^{2,3}. As a result, not all the observed spatial and seasonal patterns of atmospheric methane concentrations can be satisfactorily explained, particularly for high northern latitudes. For example, a late-autumn shoulder is consistently observed in the seasonal cycles of atmospheric methane at high-latitude sites⁴, but the sources responsible for these increased methane concentrations remain uncertain. Here we report a data set that extends hourly methane flux measurements from a high Arctic setting into the late autumn and early winter, during the onset of soil freezing. We find that emissions fall to a low steady level after the growing season but then increase significantly during the freeze-in period. The integral of emissions during the freeze-in period is approximately equal to the amount of methane emitted during the entire summer season. Three-dimensional atmospheric chemistry and transport model simulations of global atmospheric methane concentrations indicate that the observed early winter emission burst improves the agreement between the simulated seasonal cycle and atmospheric data from latitudes north of 60° N. Our findings suggest that permafrost-associated freeze-in bursts of methane emissions from tundra regions could be an important and so far unrecognized component of the seasonal distribution of methane emissions from high latitudes.

Methane emissions from permafrost dominated tundra regions are well documented^{5–7} and also recognized as considerable contributors to the dynamics of high-latitude atmospheric methane concentrations^{8,9}. The scale and dynamics of growing-season methane emissions from tundra settings have been documented mostly through flux measurements made with low time resolution using manual chambers^{5,6,10} together with some at higher time resolution taken only during the growing season^{7,11,12}. Here we report a data set that extends hourly CH₄ flux measurements from a high Arctic setting into the frozen season. The measurement site is located in Zackenberg Valley, northeast Greenland, 74.30° N 21.00° W. Six automated chambers provided flux measurements once per hour, in a typical fen area dominated by graminoids *Eriophorum scheuchzeri*, *Dupontia psilosantha* and *Arctagrostis latifolia*. Methane concentration in the chambers was measured by a laser off-axis integrated-cavity output spectroscopy analyser (Fast Methane Analyser, Los Gatos Research). The instrument sensitivity is better than 10 p.p.b.; time resolution of the primary concentration data is 1 s.

As part of the field season of the 2007 International Polar Year, the Zackenberg research station was kept open two months longer than normal. This gave us a chance to observe autumn and early-winter fluxes, which showed some surprisingly high emissions (Fig. 1;

Supplementary Table 1). This very high and variable flux happened when the active layer was gradually freezing, so CH₄ that had accumulated in this layer was probably squeezed out through the frost action. This feature has not been observed in studies at lower latitudes, possibly because the permafrost bottom is necessary to prevent CH₄ from diffusing downwards. The autumn fluxes varied greatly over small distances (chambers were less than 1 m apart), probably because peat and vegetation structure provided pathways for emission to the atmosphere. A late-autumn increase in methane emissions was observed in one of the early tundra flux studies¹³, but it lacked the time resolution needed to quantify the relative importance for the annual flux budget.

The observed growing season emission dynamics are comparable to earlier work at the same^{6,7} and at similar tundra sites¹². Integrated summer season emissions, roughly 4.5 g CH₄ m⁻² for the season, also match well with previous estimates for the same climatic and ecosystem setting^{6,7}.

Emissions decreased during September until they reached the presumed low winter emission level (Fig. 1). However, at the onset of soil freeze-in, a substantial increase in emissions was observed and was sustained for several weeks, corresponding to the time required for a complete freeze-in of the entire soil and root zone profile. Freeze-in emissions were much more variable than summer emissions. Peak emissions during the freeze-in period in individual chambers reached levels of 112.5 mg CH₄ m⁻² h⁻¹, which to our knowledge are the highest rates reported from tundra ecosystems (excluding hotspot emissions from thermokarst lakes¹⁴), and they appear at a time when previous assumptions would put tundra emissions at a negligible level (see Supplementary Information for further discussion).

Earlier studies have indicated the possibility of a spring burst from trapped methane during the winter^{15,16}. We have early-season flux data from Zackenberg for 2006 (M. Mastepanov *et al.*, manuscript in preparation) showing that spring emissions amounted to less than 2% of summer emissions (Fig. 1 insert; Supplementary Table 2), with summer emissions being very similar for 2006 and 2007 (Supplementary Tables 1 and 2). Emissions of methane during spring from this type of tundra environment are therefore not considered as a major contributor to annual methane emissions.

To investigate the potential importance of the observed methane emissions during freezing of the permafrost surface layer at large scales, we carried out model simulations of atmospheric transport and compared them with observations. Model-simulated methane concentrations were sampled at the times and locations when measurements were taken at selected background monitoring sites of the NOAA Earth System Research Laboratory's cooperative air sampling network⁴. Average seasonal cycles were constructed from air samples collected over the 4-year simulation period. Furthermore, background sites were averaged into two latitudinal bands: 25–55° N

¹GeoBiosphere Science Centre, Physical Geography and Ecosystems Analysis, Lund University, Sölvegatan 12, 22362, Lund, Sweden. ²Institute of Geography and Geology, University of Copenhagen, Øster Voldgade 10, DK-1350 Copenhagen, Denmark. ³NOAA Earth System Research Laboratory, 325 Broadway, Boulder, Colorado 80305, USA. ⁴SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands. ⁵Institute for Marine and Atmospheric Research Utrecht (IMAU), Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands. ⁶National Environmental Research Institute, University of Aarhus, Frederiksborgvej 399, 4000 Roskilde, Denmark.

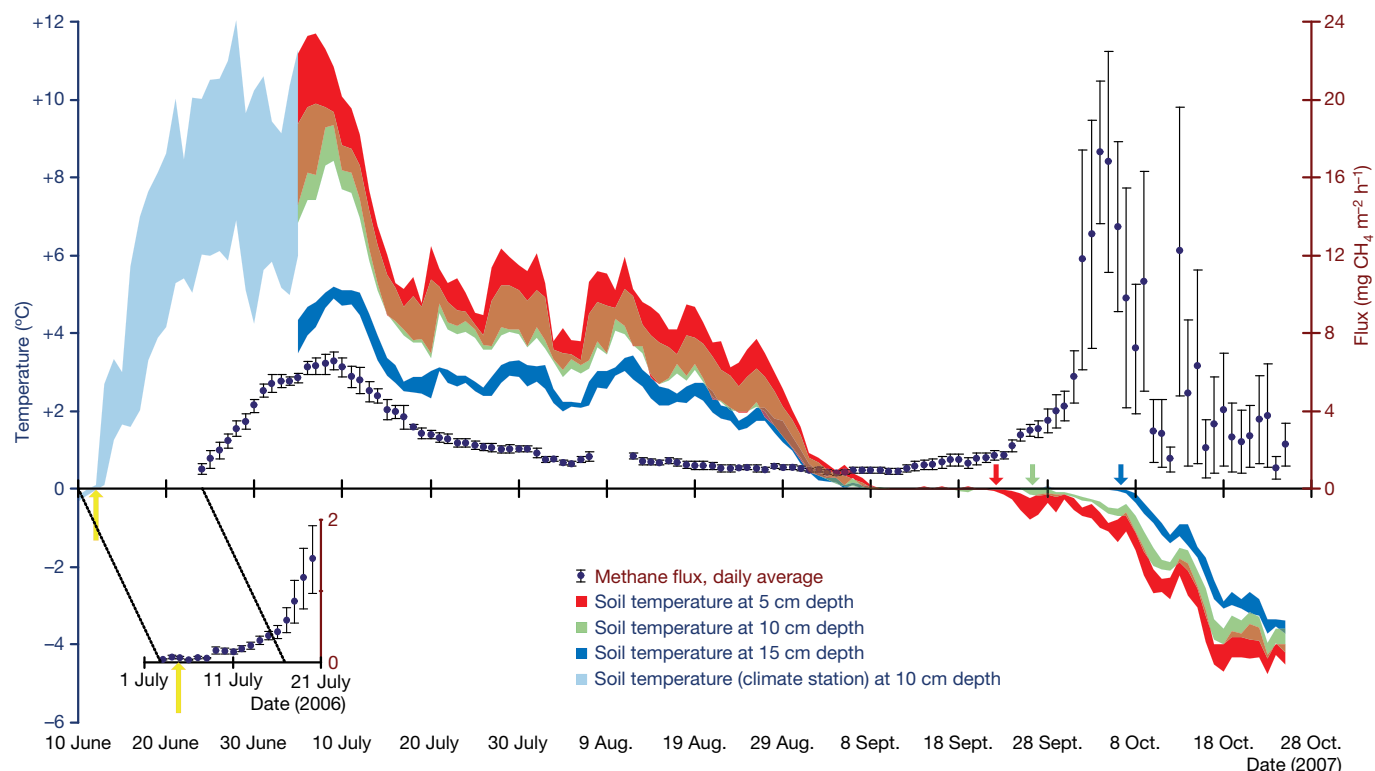


Figure 1 | Full-season methane emission and soil temperature. Soil temperatures at three depths shown as a coloured area between daily minimum and daily maximum values (5, 10 and 15 cm depth as red, green and blue). The arrows of the same colour show the date of freezing of each horizon. Soil temperatures from the nearby climate station (light blue) are shown for the period when on-site data are lacking. Site-average fluxes are shown as daily mean values averaged over six individual chambers. The error

bars show standard error of mean between the chambers. The lower inserted panel shows early-season emission in 2006 during the corresponding period relative to the date of snowmelt in 2007 (yellow arrows indicate date of snowmelt in the two years). The onset of the second emission peak coincides with freezing of the upper horizon and continues to reach a maximum when soil freezes down to -15 cm.

and 55 – 85° N (see Methods for a list of sites and where to access the CH_4 data). The averages represent 30-day running means of all samples (either modelled or measured), calculated in 5-day intervals.

There is a very reasonable agreement between the reference emission scenario (SC1) and the measurements (Fig. 2). At mid latitudes of the Northern Hemisphere, both model scenarios nicely reproduce the seasonal amplitude, although the phase lags the measurements by about a month in the second part of the year. At high northern latitudes the differences between model and measurements are more

pronounced, highlighting a deficiency of the reference model in simulating the timing of the concentration increase from summer to winter. Interestingly, the largest deviations occur in October when the unrepresented emissions from permafrost are highest. The difference between the two model simulations confirms that the influence of the simulated permafrost emissions is considerable and does improve the simulated seasonal cycle. Significant differences remain between model SC2 and the measurements, but it should be kept in mind that the underlying parameterization is only a preliminary extrapolation of the actual flux measurements. Therefore, once additional information on permafrost freeze-in emissions become available, confirmation of our model results is needed on the basis of a more sophisticated emission parameterization. Nevertheless, these results show that CH_4 emissions from the freezing active layer in permafrost areas may be an important missing process that limits model performance at high northern latitudes.

We also investigated whether there has been a change in the shape of the seasonal cycle in recent years by comparing observed seasonal cycles for the periods 1992–95 and 2002–05. The results (Supplementary Fig. 3) demonstrate that both seasonal cycles (25 – 55° N and 55 – 85° N) were remarkably constant over these periods, indicating that the signature of permafrost emissions in the observed seasonal cycle is not a recent phenomenon.

The flux measurements and atmospheric transport model results presented here are likely to be of a general nature, as there is nothing unique or artificial about this study site. It is situated in one of the most pristine environments in the world (the National Park of north-east Greenland) and there is no reason why such a physical mechanism should not happen everywhere that there are similar ecosystems. This study benefited from the unique opportunity through the International Polar Year effort to keep the Zackenberg

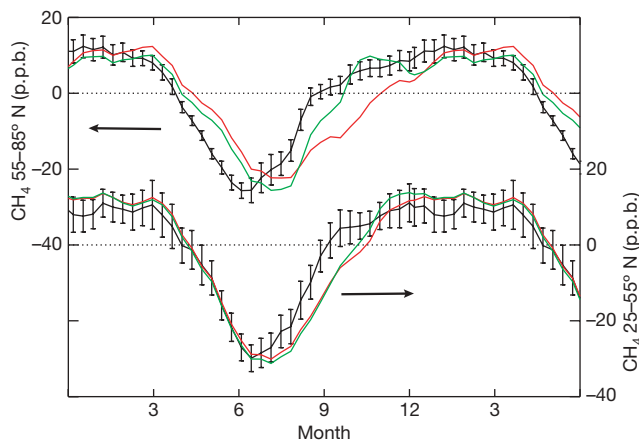


Figure 2 | Comparison of measured and model-simulated latitudinally averaged seasonal cycles of methane. Black, measurements with 2 sigma uncertainty intervals; red, the model simulation using the reference scenario; green, the model simulation including a representation of additional emissions from freezing permafrost (see text).

station open for longer than usual, and thus to observe a phenomenon that has most likely been missed in other measurements around the circumpolar north because of the difficulties of maintaining flux measurements into the frozen season at remote high-emitting wet tundra sites. If the fluxes measured at Zackenberg are applied to all of $0.88 \times 10^{12} \text{ m}^2$ of wet meadow tundra¹⁷ (disregarding possible similar emissions from mesic tundra which covers even greater areas), it will amount to a pulse of $\sim 4 \text{ Tg CH}_4$ from the highest latitudes at what was previously thought to be an inactive time of year in terrestrial ecosystems. This is in agreement with a corresponding estimate based on the three-dimensional modelling which amounts to 3.9 Tg CH_4 (see Methods). This does not greatly increase emission estimates from high northern latitudes, but it revises our view of the seasonal distribution of known emissions.

METHODS SUMMARY

Methane emissions were measured by an automatic chamber method; flux was calculated from the increase in the chamber CH_4 concentration, corrected by air temperature and pressure. Erroneous measurements (for instance, during strong southerly winds that tend to cause improper closing of the chambers) were filtered out and no artificial corrections or gap filling were applied.

Global methane concentrations were simulated using an atmospheric chemistry and transport model, which includes a dedicated representation of the methane cycle. Calculations were performed with and without a parameterization of methane emissions from freezing permafrost. Simulated concentrations were compared with high precision methane measurements representative of the background conditions at mid to high northern latitudes.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 April; accepted 18 September 2008.

1. Mikaloff Fletcher, S. E., Tans, P. P., Bruhwiler, L. M., Miller, J. B. & Heimann, M. CH_4 sources estimated from atmospheric observations of CH_4 and its $^{13}\text{C}/^{12}\text{C}$ isotopic ratios: 1. Inverse modeling of source processes. *Glob. Biogeochem. Cycles* **18**, GB4004 (2004).
2. Dlugokencky, E. J. *et al.* Atmospheric methane levels off: Temporary pause or a new steady-state? *Geophys. Res. Lett.* **30**, 1992 (2003).
3. Miller, J. B. *et al.* Airborne measurements indicate large methane emissions from the eastern Amazon basin. *Geophys. Res. Lett.* **34**, L10809 (2007).
4. Dlugokencky, E. J., Steele, L. P., Lang, P. M. & Masarie, K. A. The growth rate and distribution of atmospheric methane. *J. Geophys. Res.* **99**, 17021–17043 (1994).
5. Reeburgh, W. S. *et al.* A CH_4 emission estimate for the Kuparuk River basin, Alaska. *J. Geophys. Res.* **103**, 29005–29013 (1998).
6. Christensen, T. R. *et al.* Trace gas exchange in a high-arctic valley 1. Variations in CO_2 and CH_4 flux between tundra vegetation types. *Glob. Biogeochem. Cycles* **14**, 701–713 (2000).
7. Friborg, T., Christensen, T. R., Hansen, B. U., Nordstroem, C. & Soegaard, H. Trace gas exchange in a high-arctic valley 2. Landscape CH_4 fluxes measured and modeled using eddy correlation data. *Glob. Biogeochem. Cycles* **14**, 715–723 (2000).
8. Gedney, N., Cox, P. M. & Huntingford, C. Climate feedback from wetland methane emissions. *Geophys. Res. Lett.* **31**, L20503 (2004).
9. Bousquet, P. *et al.* Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature* **443**, 439–443 (2006).
10. Christensen, T. R. Methane emission from Arctic tundra. *Biogeochemistry* **21**, 117–139 (1993).
11. Fan, S. M. *et al.* Micrometeorological measurements of CH_4 and CO_2 exchange between the atmosphere and subarctic tundra. *J. Geophys. Res.* **97**, 16627–16644 (1992).
12. Corradi, C., Kolle, O., Walter, K., Zimov, S. A. & Schulze, E.-D. Carbon dioxide and methane exchange of a north-east Siberian tussock tundra. *Glob. Change Biol.* **11**, 1910–1925 (2005).
13. Whalen, S. C. & Reeburgh, W. S. A methane flux time series for tundra environments. *Glob. Biogeochem. Cycles* **2**, 399–409 (1988).
14. Walter, K. M., Zimov, S. A., Chanton, J. P., Verbyla, D. & Chapin, F. S. III. Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature* **443**, 71–75 (2006).
15. Hargreaves, K. J., Fowler, D., Pitcairn, C. E. R. & Aurela, M. Annual methane emission from Finnish mires estimated from eddy covariance campaign measurements. *Theor. Appl. Climatol.* **70**, 203–213 (2001).
16. Tokida, T. *et al.* Episodic release of methane bubbles from peatland during spring thaw. *Chemosphere* **70**, 165–171 (2007).
17. Bliss, L. C. & Matveyeva, N. V. in *Arctic Ecosystems in a Changing Climate: An Ecophysiological Perspective* (eds Chapin, F. S. III, Jefferies, R. L., Reynolds, J. F., Shaver, G. R. & Svoboda, J.) 59–89 (Academic, 1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was carried out under the auspices of the GeoBasis programme and part of the Zackenberg Ecological Research Operations funded by the Danish Ministry of the Environment and the ISICaB project funded by the Commission for Scientific Research in Greenland (KVUG). ASIAQ–Greenland Survey provided climate data. The work was also supported by the Swedish Research Councils VR and FORMAS. We thank P. Bergamachi (JRC) and J.-F. Meirink (KNMI) for providing the TM5 model setup. T. Tagesson helped with the field work in Zackenberg. We are grateful for comments on earlier versions of this manuscript from A. Lindroth and B. Christensen.

Author Contributions T.R.C., M.P.T., M.M., C.S. and L.S. designed the field research; M.M. designed, constructed and set up the automatic measurement system in Zackenberg; C.S. operated the system and performed manual measurements; M.M. performed data analysis; E.D. and S.H. provided atmospheric CH_4 data and designed and ran the atmospheric transport model experiments; T.R.C., M.M., S.H. and E.D. drafted the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.R.C. (Torben.Christensen@nateko.lu.se).

METHODS

Zackenberg site description. The Zackenberg Valley is situated at 74° 30' N, 21° 00' W in the National Park of northeast Greenland. A research station (Zackenberg Research Station) was established in 1997 and offers basic logistical facilities (an airstrip, laboratories, satellite-based communication systems and so forth) necessary for carrying out efficient research. The site has a short record of meteorological observations, and 1996 was the first full year when basic meteorological variables were registered continuously. Mean annual temperature in the first 10-year period of the station ranged from -8.5°C to -10.1°C with July as the warmest month (mean monthly air temperature of 5.8°C) and February as the coldest month (mean monthly air temperature of -22.4°C). The average frost-free period during these 10 years was 35 days, lasting from mid July to late August. In Daneborg, situated on the outer coast 22 km southeast of Zackenberg and with a longer period of meteorological measurements, the mean annual temperature for the period 1960–90 was -10.3°C . The warmest month, July, had a mean of 3.8°C and February, the coldest month, had a monthly mean of -17.6°C . The valley is dominated by minerotrophic sedge-grass-rich fens mixed with elevated areas of dwarf shrub heaths with *Cassiope tetragona* and *Salix arctica* as dominant species. A slightly sloping fen area is the main study area here. The peat layer in the fen is 20–30 cm thick, typical of high arctic fen ecosystems¹⁸. Onset of peat accumulation has been ^{14}C -dated to AD 1290–1390 in a neighbouring fen area, and the surrounding Little Ice Age nival fans and nivation basins primarily contain organic material deposited from AD 1420 to 1500–1580¹⁹.

The active layer depth specifically on the measurement site reached 50–56 cm (near different chambers) before soil freezing in 2007. Despite the low temperatures, the snow cover was mosaic until 20 October, and then was no more than 3 cm deep until the first snowstorm on 26 October.

A large body of background information from the Zackenberg Research Station has recently been summarized in a book volume celebrating the first 10 years of activities at the research station¹⁸.

Methane flux measurements. Automatic chambers were deployed in August 2005 and the first seasonal data set was obtained in 2006 (3 July to 26 August). In 2007 an extended season was carried out (26 June to 25 October). Six chambers have been aligned in a row from the periphery of the fen towards its central part. The distance between the chambers is 30–80 cm. The chambers are made of Plexiglas with aluminium corners. Each chamber is 60 × 60 cm and about 30 cm height (depending on microtopography). The chamber lid stays open for 55 min per hour and closes for five minutes for the measurements. Air is mixed in a closed chamber by a fan; the same fan ventilates a chamber when it is open. Air from the chamber passes through 30 m of tubing (internal diameter 4 mm) to the analytical box and after the non-destructive analysis it goes back to the chamber. The analytical box contains a methane analyser (Fast Methane Analyser, Los Gatos Research), CO_2 analyser (SBA-4, PP Systems) and solenoid valves. The concentration data are collected at 1 Hz rate; data acquisition starts three minutes before a chamber closes, continues for five minutes while it is closed, and then two minutes after the chamber opens, so the full cycle of six chambers takes one hour.

Although we did not make direct measurements of soil temperature and humidity inside the chambers, to avoid extra disturbance, the visual control does not give any evidence of the construction affecting the temperature and water regime inside the chambers. Visible water table and the snow level (during the snowfall and snowmelt) is the same inside and outside the chambers.

The CH_4 fluxes are calculated from the slope of concentration change in the closed chamber; if the increase was not linear during five minutes of closure, the most linear part of this time is arbitrarily chosen. The air temperature and pressure for flux calculations are obtained from Zackenberg micrometeorological station located about 1 km from the site. We made additional measurements of water table level, active layer depth, PAR, soil temperature and humidity next to the chambers.

Atmospheric CH_4 measurements and chemical transport model. Atmospheric CH_4 measurements are from weekly samples collected at sites in the NOAA Earth System Research Laboratory's cooperative global air sampling network⁴. We determined methane dry-air mole fractions by gas chromatography with flame ionization detection against the WMO CH_4 mole fraction standard scale. Over the period of this study, repeatability of the measurements (1σ) was ~ 2 p.p.b. Latitudinal averages contained the following sites: 25–55° N contained 'mid', 'bme', 'bmw', 'uta', 'mhd', 'ask', 'nwr', 'pta', 'azr', 'izo', 'pocn30' and 'pocn25', and 55–85° N contained 'alt', 'shm', 'brw', 'ice', 'zep', 'stm', 'cba' and 'sum' (see www.esrl.noaa.gov/gmd/ccgg/flask.html for a list of site codes and [ftp://ftp.cmdl.noaa.gov/ccg/ch4/flask/event](http://ftp.cmdl.noaa.gov/ccg/ch4/flask/event) for access to data).

We made atmospheric transport model calculations using the TM5 model²⁰ for the period 2002–05, at a spatial resolution of $6^{\circ} \times 4^{\circ}$ and 25 vertical sigma-pressure levels. Two scenarios of methane sources and sinks were applied: a reference scenario (SC1) and a scenario including emissions from permafrost freeze-in (SC2). The methane sources and sinks of SC1 correspond with the a priori assumptions that were used in the inverse modelling calculations of ref. 21, with the exception of wetlands. Wetland emissions were taken from ref. 22, and rescaled to a global total of 175 Tg $\text{CH}_4 \text{ yr}^{-1}$ and high-latitude (50–90° N) emissions of 20 Tg $\text{CH}_4 \text{ yr}^{-1}$. SC2 is the same as SC1 except for additional emissions from freezing permafrost. For lack of any detailed information on this process, we followed a highly simplified procedure, assuming that emission started when the diurnal mean temperature dropped below -2°C and continued for a period of 1 month. This process was only active in those model grid boxes that were classified as continuous or discontinuous permafrost according to the CAPS circumpolar permafrost map²³. The annual emission of freezing permafrost was assumed to be the same as the (summer time) wetland emission in each model grid box for which the process is active. This procedure introduces an additional source of 3.9 Tg $\text{CH}_4 \text{ yr}^{-1}$, which moves from north to south during autumn and reaches maximum global emissions in October.

18. Meltofte, H., Christensen, T. R., Elberling, B., Forchhammer, M. C. & Rasch, M. (eds) *High-Arctic Ecosystem Dynamics in a Changing Climate. Advances in Ecological Research* Vol. 40 (Elsevier, 2008).
19. Christiansen, H. H. *et al.* Holocene environmental reconstruction from deltaic deposits in northeast Greenland. *J. Quat. Sci.* **17**, 145–160 (2002).
20. Krol, M. C. *et al.* The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmos. Chem. Phys.* **5**, 417–432 (2005).
21. Bergamaschi, P. *et al.* Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT. 2. Evaluation based on inverse model simulations. *J. Geophys. Res.* **112**, D02304 (2007).
22. Walter, B. P., Heimann, M. & Matthews, E. Modeling modern methane emissions from natural wetlands. 2. Interannual variations 1982–1993. *J. Geophys. Res.* **106**, 34207–34217 (2001).
23. Brown, J., Ferrians, O. J. Jr, Heginbottom, J. A. & Melnikov, E. S. *Circum-arctic Map of Permafrost and Ground-Ice Conditions*. USGS Circum-Pacific Map Series CP-45 (scale 1:10,000 000) (US Geological Survey, 1997).

Partial rupture of a locked patch of the Sumatra megathrust during the 2007 earthquake sequence

A. Ozgun Konca¹, Jean-Philippe Avouac¹, Anthony Sladen¹, Aron J. Meltzner¹, Kerry Sieh^{1†}, Peng Fang², Zhenhong Li³, John Galetzka¹, Jeff Genrich¹, Mohamed Chlieh⁴, Danny H. Natawidjaja¹, Yehuda Bock², Eric J. Fielding⁵, Chen Ji⁶ & Don V. Helmberger¹

The great Sumatra–Andaman earthquake and tsunami of 2004 was a dramatic reminder of the importance of understanding the seismic and tsunami hazards of subduction zones^{1–4}. In March 2005, the Sunda megathrust ruptured again, producing an event⁵ of moment magnitude (M_w) 8.6 south of the 2004 rupture area, which was the site of a similar event in 1861 (ref. 6). Concern was then focused on the Mentawai area, where large earthquakes had occurred in 1797 ($M_w = 8.8$) and 1833 ($M_w = 9.0$)^{6,7}. Two earthquakes, one of $M_w = 8.4$ and, twelve hours later, one of $M_w = 7.9$, indeed occurred there on 12 September 2007. Here we show that these earthquakes ruptured only a fraction of the area ruptured in 1833 and consist of distinct asperities within a patch of the megathrust that had remained locked in the interseismic period. This indicates that the same portion of a megathrust can rupture in different patterns depending on whether asperities break as isolated seismic events or cooperate to produce a larger rupture. This variability probably arises from the influence of non-permanent barriers, zones with locally lower pre-stress due to the past earthquakes. The stress state of the portion of the Sunda megathrust that had ruptured in 1833 and 1797 was probably not adequate for the development of a single large rupture in 2007. The moment released in 2007 amounts to only a fraction both of that released in 1833 and of the deficit of moment that had accumulated as a result of interseismic strain since 1833. The potential for a large megathrust event in the Mentawai area thus remains large.

Slip along a subduction megathrust can be either aseismic or seismic. Seismic slip commonly has a duration of seconds to minutes, with sliding velocities of about a metre per second and rupture velocities of a few kilometres per second⁸. Such rapid failure generates strong ground shaking and tsunamis. Slower, aseismic slip is also common, and dominates at depths greater than about 40 km (ref. 9) but also occurs at shallower depths^{10–14}. This process leads to heterogeneous strain accumulation in the interseismic period, with stress building up around locked patches that presumably fail during megathrust earthquakes.

The modelling of geodetic and palaeogeodetic measurements of interseismic strain indeed shows that the Sunda megathrust, off the shore of Sumatra, consists of a patchwork of creeping and locked areas¹⁵ (Fig. 1) and suggests some correlation between megathrust earthquakes and interseismic coupling. A recent example of this is the rupture of a 350-km section, in the 2005, $M_w = 8.6$ Nias earthquake⁵. Historical accounts⁶ and palaeoseismic data¹⁶ indicate that rupture of almost the same patch produced an earlier great earthquake in 1861. Immediately south of the Nias patch, near the Equator, coupling is

low and consistent with there having been only moderate earthquakes in the past few centuries^{10,17}. Farther south, beneath the Mentawai Islands, coupling has been high for at least the past 40 years and great earthquakes have occurred there¹⁸. These observations indicate that the pattern of interseismic strain accumulation has a strong influence on the characteristics of large megathrust ruptures.

Nonetheless, successive seismic ruptures in the same area may differ significantly in extent and magnitude^{19–21}. This Letter provides a very well-documented case of such variability. We show that the 2007 sequence of large seismic ruptures on the strongly coupled Mentawai patch differs significantly from previous ruptures of this section in 1833 and 1797. The slip and extent of the recent ruptures are much smaller than during those previous events. Moreover, the amount of moment released during the recent events is much less than the deficit of moment that has accumulated since the previous great earthquakes.

We use Global Positioning System (GPS) measurements, field measurements of uplift, synthetic aperture radar (SAR) interferometry (InSAR) and seismological records to estimate the source parameters of the two large earthquakes of 12 September 2007. Details are given in Supplementary Information. The 2007 ruptures occurred beneath a large subset of the continuously recording GPS (cGPS) stations of the Sumatran GPS array. Displacements assigned to each of the events were determined from 120-s time series. By contrast, the InSAR and field measurements provide information only on the cumulative effects of the sequence of earthquakes. These measurements cover greater lengths of time and therefore must include some amount of postseismic deformation as well.

All horizontal GPS motions are trenchwards (Fig. 2a). The maximum horizontal displacement due to the cumulative effect of the whole sequence, 1.5 m, occurred at station BSAT on South Pagai Island. Vertical displacements were measured at the GPS stations and from the emergence or submergence of coral microatolls on the reefs that fringe the Mentawai Islands. They show uplift of the islands and subsidence of the mainland coast (Fig. 2a, inset). The maximum uplift reaches 1.3 m on Mega Island, about 70 km northwest of the epicentre. Uplift decreases northwards to about 1 m on southern South Pagai Island to 10 cm on North Pagai Island. Uplift at Sipora Island is on the order of 20 cm. These data are complemented with InSAR line-of-sight displacements. Using ROI_PAC (repeat-orbit interferometry package) software²², we processed four independent interferograms from phased-array-type L-band SAR (PALSAR) images acquired by the Advanced Land Observing Satellite (Japan Aerospace Exploration Agency). These data are consistent with the

¹Tectonics Observatory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA. ²Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA. ³COMET, Department of Geographical and Earth Sciences, University of Glasgow, Glasgow G12 8QQ, UK.

⁴Geosciences Azur, Université de Nice Sophia-Antipolis, Institut de Recherche pour le Développement, Observatoire de la Côte d'Azur, 06560 Valbonne, France. ⁵Jet Propulsion Laboratory, M.S. 300-233, California Institute of Technology, Pasadena, California 91109, USA. ⁶Department of Geological Sciences, University of California, Santa Barbara, California 94106. [†]Present address: Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Avenue, 639798 Singapore.

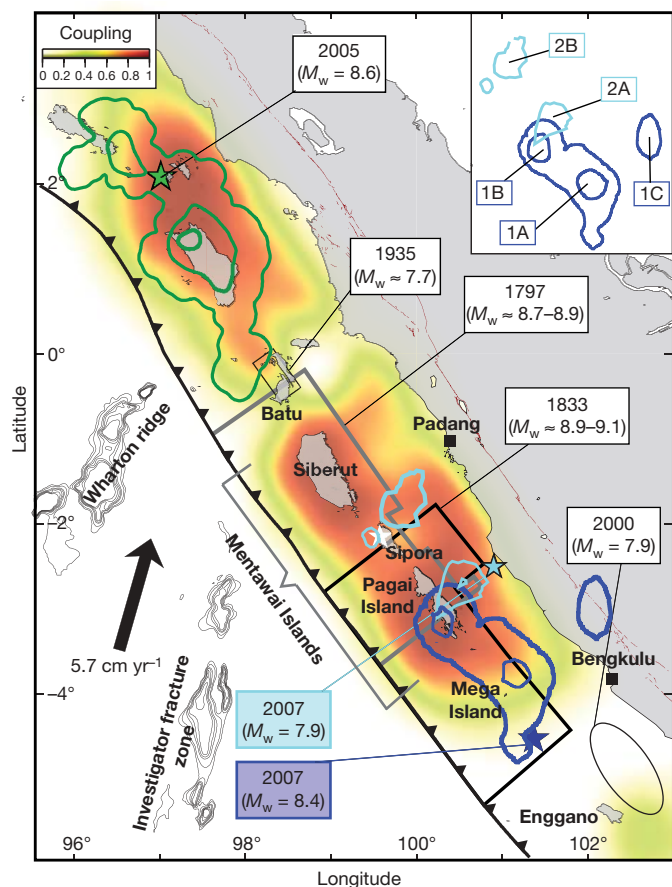


Figure 1 | Patches with strong interseismic coupling on the Sunda megathrust coincide with large seismic ruptures. The pattern of coupling, defined as the ratio of interseismic slip rate to plate convergence rate, is derived from the modelling of geodetic and paleogeodetic data¹⁵. No information is available on coupling under northern Simeulue, west of about 96.2° E. The black arrow shows the 5.7 cm yr⁻¹ plate convergence rate. Slip distribution of the $M_w = 8.6$ earthquake of 2005 is shown using 5-m contour lines in green⁵. Gray and black polygons show estimated rupture areas of the 1797 and 1833 earthquakes⁷. Dark and pale blue lines show the 1-m and 5-m slip contour lines of the $M_w = 8.4$ and $M_w = 7.9$ seismic ruptures of 2007, respectively; stars show the epicentres. The smaller, $M_w = 7.7$, earthquake of 1935 (ref. 17) occurred in a region of weak coupling. The $M_w = 7.9$ earthquake of 2000, which is predominantly an intraslab strike-slip event³⁰ also falls in an area with low coupling. The inset map displays various asperities of the 2007 rupture sequence.

GPS and coral measurements where they coincide, and provide a much denser spatial coverage. They show a strong displacement gradient under the Pagai Islands and reveal a zone of deformation beneath an area north of Bengkulu on mainland Sumatra (Supplementary Fig. 1).

We determine static and kinematic source models from the modelling of the surface displacements and teleseismic waveforms recorded at a selection of IRIS (Incorporated Research Institutions for Seismology) stations chosen to ensure good azimuthal coverage. The geometry of the megathrust is approximated by a plane dipping 15° to the northeast, away from the trench. We also derive models that feature a curved fault geometry with the dip angle increasing with depth (Supplementary Information). These tests show that the results described here are independent of the assumed geometry of the megathrust. We first derived a cumulative slip model that includes the $M_w = 8.4$ and $M_w = 7.9$ earthquakes using the cGPS cumulative displacements, the field measurements of uplift and the InSAR data. The model obtained using only the cGPS displacements measured from just before the $M_w = 8.4$ event until just after the $M_w = 7.9$ event is the least contaminated by postseismic relaxation. It suggests that there is a relatively patchy slip distribution with a

geodetic moment of 7.3×10^{21} N m (Supplementary Fig. 2a). The spatial resolution of the model improves when the InSAR and field data are added, but some contamination by postseismic deformation is introduced. The best-fitting model calculated from all these data (Fig. 2a) has a total moment of 7.5×10^{21} N m (equivalent to $M_w = 8.5$), which is only marginally larger than that derived from the cGPS measurements alone. Thus, we consider this source model to be a better-constrained representation of the coseismic slip distribution than the model derived from only GPS data. This cumulative source model is also very similar to that obtained by summing the $M_w = 8.4$ and $M_w = 7.9$ models and a $M_w = 7.0$ aftershock model (Supplementary Fig. 2b). The greater resolution afforded by the coral and InSAR measurements shows that the patchiness of the slip distribution is real and not an artefact of our methodology.

The cumulative slip model (Fig. 2a) has a dumb-bell-shaped principal rupture area extending contiguously northwestwards from the hypocentre, off the shore of Bengkulu, to South Pagai Island. In addition, there are two disconnected minor slip patches. One is below northern Sipora Island, about 100 km along the strike direction of the northwestern edge of the principal rupture. The other lies beneath the volcanic arc of the mainland, about 100 km down-dip from the down-dip edge of the principal rupture. Slip peaks at 8 m on the main patch under southern South Pagai Island and at 5-m local maxima about 25 km northeast of Mega Island. The maximum slip on the small patch beneath Sipora Island is 2.5 m. All data sets are fitted well by slip on a single plane representing the megathrust; slip on any other fault is therefore not required.

We also derived separate kinematic source models of the mainshock and principal aftershock using teleseismic waveforms, GPS measurements and subsets of the coral and InSAR data. For this analysis, we discarded data from the Pagai Islands, where the contributions from each of the two events cannot be distinguished. Farther south and east, the displacements measured along PALSAR track 445 and coral measurements on Mega Island are clearly attributable to the mainshock alone (Supplementary Information, section H). We modelled these subsets of the coral and InSAR data together with the GPS measurements and the teleseismic records of the mainshock.

The source model of the $M_w = 8.4$ mainshock shows unilateral northward rupture, which started about 70 km south of Mega Island (Fig. 2b). As in the cumulative source model, the most prominent slip loci are under southern South Pagai Island, where slip peaks at about 7 m; about 25 km north of Mega Island, where slip peaks at about 6 m; and on the deep patch east of the Sumatran coast (Fig. 1, inset). The total seismic moment of this model ($\sim 5 \times 10^{21}$ N m) is consistent with the Global Centroid Moment Tensor moment magnitude (<http://www.globalcmt.org/>). The source time function indicates that the rupture was not very impulsive and lasted for about 100 s (Fig. 2b, inset). The rise times are estimated to be 5–10 s, and the moment rate increased smoothly over the first 20 s (Supplementary Fig. 5). The southern patch slipped during the first 40 s and the northern patch slipped during the second 40 s.

The model of the $M_w = 7.9$ earthquake derived from the joint inversion of the teleseismic and cGPS data shows that this earthquake also involved failure of more than one discrete patch (Fig. 2c). The moment of this event, 1.1×10^{21} N m, was released in two pulses over about 80 s. (Fig. 2c, inset). Rupture began within a few kilometres of the down-dip edge of the mainshock's northern patch. The seismic waveforms require an extremely abrupt initiation of the first subevent (rise times of a few seconds at most) and a highly peaked slip distribution around the hypocentre. The second subevent occurred in 50–80 s and about 130 km farther northwest of the epicentre, east of northeastern Sipora Island. There is no evidence for significant slip between these two subevents.

The September 2007 sequence ruptured a number of distinct asperities (defined here as patches with locally large slip) on the megathrust that lie both within a patch that had remained strongly locked in the interseismic period and within the rupture area of the 1833

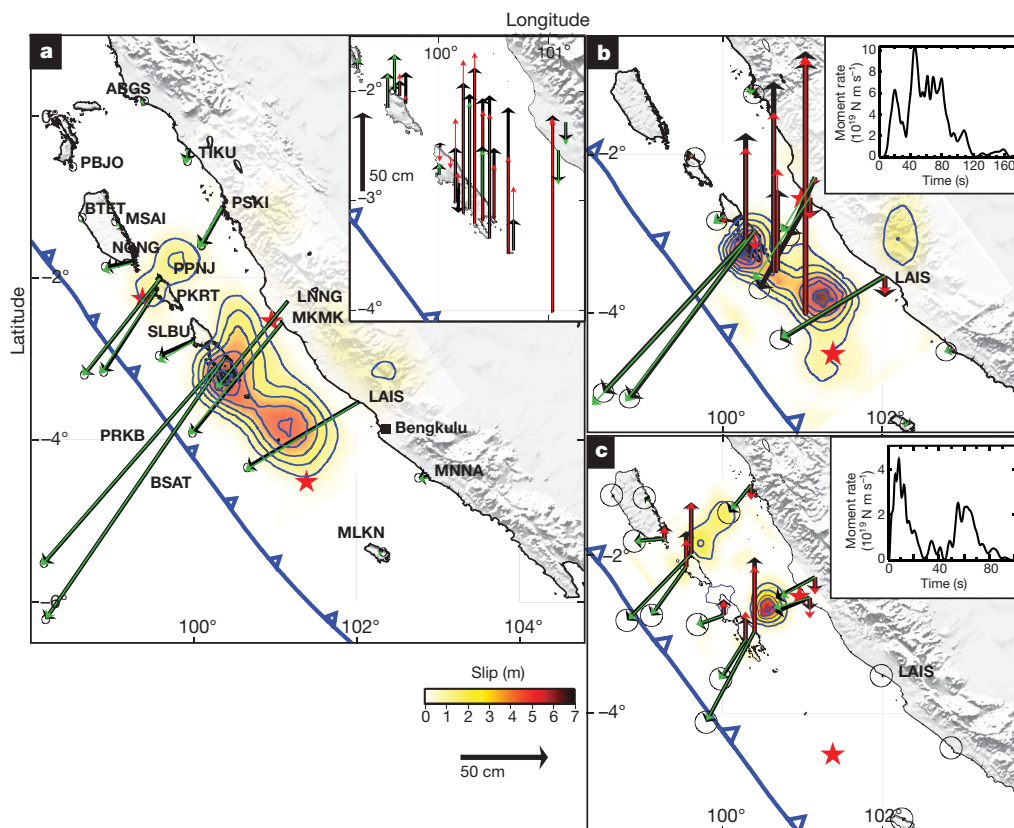


Figure 2 | Models of megathrust slip during the $M_w = 8.4$ and $M_w = 7.9$ earthquakes show principal slip on widely separated patches. All slip models and GPS vectors are plotted on the scale represented by the 50-cm arrow. Slip contour lines are plotted every 1 m, starting at 1 m. **a**, Cumulative slip distribution due to the $M_w = 8.4$ and $M_w = 7.9$ earthquakes of 12 September 2007. Contours show cumulative slip of best-fitting inversion of the GPS, coral and InSAR data. Vectors show observed (black) and modelled (green) horizontal displacement values at cGPS stations of the Sumatran GPS array (<http://www.tectonics.caltech.edu>). Normalized error ellipses are also shown. Inset, vertical GPS displacements and measurements of coral uplift (black) and the fits from the model (GPS, green; coral, red). **b**, Slip model of the $M_w = 8.4$ earthquake obtained from the joint inversion of teleseismic waveforms, GPS data, InSAR data and measurements of coral

uplift unambiguously attributable to the $M_w = 8.4$ event. Observed displacements, shown with error ellipses in black, originate at the GPS station locations and coral sites. Modelled horizontal and vertical vectors are green and red, respectively. Red stars represent the locations of epicentre as accepted by the US Geological Survey. Inset, moment rate function. Geodetic data are tabulated and fits to the InSAR and teleseismic data are shown in Supplementary Information. **c**, Model of the $M_w = 7.9$ earthquake from the joint inversion of teleseismic waveforms and GPS measurements. Vector colour coding is the same as in **b**. Inset, the moment rate function shows that the moment was released in two discrete episodes, about 20 s apart. Fits to the teleseismic data are shown in the Supplementary Information.

earthquake (Fig. 1). However, the patterns and amounts of slip in 1833 and 2007 are significantly different. Coseismic uplifts in 1833 (between 1 and 2.5 m from South Pagai Island to Sipora Island)⁷, are much larger than those observed in 2007. This is consistent with the cumulative geodetic moment of 7.5×10^{21} N m released by the 2007 earthquake sequence, representing a fraction of the moment of $(10\text{--}55) \times 10^{21}$ N m released in 1833 (ref. 15; Fig. 3). The coast of North Pagai Island was uplifted by 2.2 m in 1833 (ref. 7). This area is clearly a low-slip patch in 2007, as indicated by the modest horizontal and vertical displacements recorded at station SLBU (22 cm and 7 cm, respectively). Thus, it acted more like a 'barrier' during the coseismic slip in 2007.

South of latitude 2° S, the moment deficit accumulated since 1833 is still less than the moment released during the 1797 and 1833 events (Fig. 3). North of this latitude, the accumulated deficit is far greater than the moment released during the 1797 and 1833 events. Thus, the next great rupture might have been expected to occur north of 2° S; instead, the 2007 events occurred south of 2° S. Furthermore, the moment released during the 2007 sequence is much less than that released during the 1833 event and much less than that which has accumulated since then. These relationships clearly demonstrate that the Mentawai patch is behaving in neither a slip- nor a time-predictable manner. If rupture were time predictable, slip would already have

occurred north of 2° S. If rupture were slip predictable, slip would have been far greater in 2007 south of 2° S.

Two salient questions are why the 2007 sequence did not duplicate the 1833 event and why it released only about 25% of the deficit of moment that had accumulated since 1833. The 2007 sequence consisted of several spatially and temporally separate asperities that probably did not cooperate effectively. If two neighbouring asperities on the same fault plane rupture jointly, they are expected to cooperate to release more moment than if they had ruptured independently²³. This is because the static stress change induced by one asperity increases both the stress on the other and, hence, the elastic stress that is released during the rupture.

Spatio-temporal evolutions of the 2007 ruptures show that this kind of cooperation was not effective. For example, slip on the second asperity of the $M_w = 7.9$ earthquake (2B in Fig. 1; see inset for all labelled asperities) started only once the slip on the first asperity (2A) was over. The reloading of asperity 2A due to rupture of asperity 2B thus did not contribute to any additional slip. This may be because the area between the two asperities acted as a barrier to rupture propagation. This intervening area beneath North Pagai Island experienced little coseismic slip in 2007, but is probably not a permanent barrier, because the same area is predicted to have experienced the largest cumulative slip (of about 17 m) by the sum of the

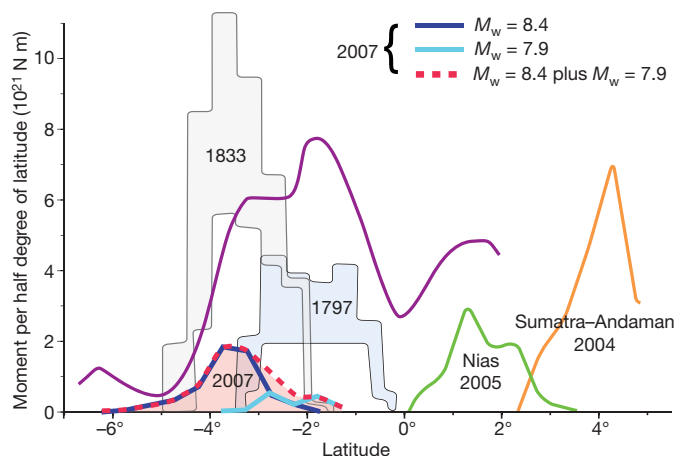


Figure 3 | Comparison of the moment released in 2007 with the moment deficit accumulated since the 1797 and 1833 earthquakes. Light grey curves are the upper and lower bounds of estimated moment release for the 1833 and 1797 earthquakes, based on modelling of geodetic and paleogeodetic data¹⁵ and taking into account the effect of 20% postseismic slip. The confidence intervals were deduced from the uncertainties in the extent of the ruptures and in the eventual contribution of postseismic deformation to the vertical displacements deduced from the coral data. The purple curve shows accumulated moment deficit since the last great rupture, derived from the modelling of interseismic strain accumulation (Fig. 1, model J-a of ref. 15). Moment deficit values are integrals within bins that are one-half degree of latitude wide.

slip models of the 1797 and 1833 earthquakes⁷. This area may therefore have acted as a barrier in 2007 because it had a locally lower stress level before the earthquake, as a result of release during previous earthquakes. Minimal cooperation between the rupture of asperities 1B and 2A is also evident from the 12-h time lag between their ruptures. The cause of the lack of cooperation between these two asperities is more enigmatic, given that they lie so close to each other. A narrow zone with low pre-stress due to the slip distribution related to the 1833 and 1797 events may have acted as a barrier, or there may be a creeping zone too narrow to be seen in the pattern of interseismic strain. In any case, it appears that the static stress increase on asperity 2A due to the $M_w = 8.4$ earthquake was enough to trigger a delayed rupture of this asperity. By contrast, the dynamic stresses induced by the $M_w = 8.4$ earthquake failed to trigger the rupture of asperity 2A.

Two independent SAR interferograms and the GPS displacement at station LAIS show that the $M_w = 8.4$ rupture induced a localized surface deformation just north of Bengkulu, and that this deformation took place during the earthquake, even though it did not contribute much to the seismic radiation (Supplementary Information). It is possible to model this signal as a deep slip patch on the megathrust that falls in a zone that creeps in the interseismic period (asperity 1C). Because this slip patch is isolated, we exclude the possibility that it is a result of the rupture propagating into the rate-strengthening zones. It could reflect seismic rupture of a rate-weakening portion of the megathrust embedded in a dominantly creeping zone, or it may be an example of a triggered aseismic transient. Deep aseismic transients on megathrusts have been observed¹⁴ and justified on the basis of rate and state friction theory^{24,25}. Another possibility is that this deformation did not take place on the megathrust but at lesser depths. The available data do not resolve this ambiguity.

In conclusion, the rupture area of the 2007 Mentawai earthquakes was confined to a subset of a locked portion that is surrounded by creep during the interseismic period. Such permanent barriers, which are found to influence the down-dip extent as well as the lateral extent of megathrust ruptures, can be imaged from the modelling of interseismic strain^{2,12–15} except when they lie in stress shadows, in particular along the up-dip portion of the plate interface¹¹. The complex

spatio-temporal pattern of the 2007 rupture is probably related to the fact that it produced much less slip than did historical earthquakes in the area. The 2007 ruptures released only 25% of the deficit of moment that had accumulated since the last rupture. The sequence essentially ruptured a set of asperities, which triggered each other through static and dynamic interactions but did not cooperate efficiently, because of the intervening barriers. Some of these barriers are most likely not permanent and are related to the slip in past earthquakes. Whereas permanently creeping barriers should tend to favour some regularity and similarity of earthquakes, the presence of non-permanent barriers due to the stress distribution left over from previous ruptures is probably the major factor introducing irregularity, as observed in dynamic fault models^{26,27}. This is probably the main reason that neither the slip-predictable nor the time-predictable models apply, and why the 2007 earthquakes did not grow as big as the 1833 earthquake. This supports the view that seismic asperities are probably not permanent features²⁸ but rather move from one rupture to another within the area that is locked in the interseismic period.

METHODS SUMMARY

We assume that the seismic rupture can be modelled as the propagation of a rupture front with finite width. The kinematic source model is then characterized by the static slip, the rupture velocity and the rise time, namely the time it takes for slip at a point on the fault to reach its final static displacement, at each elementary patch. The joint inversion consists of searching for the model that best fits the wavelet transform of the seismograms, the geodetic data and the InSAR data. We use an optimization method based on a simulated annealing algorithm, where bounded parameter spaces of slip amplitude, rake angle and rupture velocity are searched to obtain models that fit both teleseismic and geodetic data²⁹. Rupture velocity is allowed to vary between 2.1 and 2.8 km s⁻¹, and the rake angle can vary between 80° and 130°. We selected teleseismic waveforms from the IRIS network to ensure good azimuthal coverage, constraining the source model. The broadband seismograms were band-pass filtered from 1.5 s (P waves) and 3 s (SH waves) to 200 s. We used 16 P and 19 SH waveforms for the $M_w = 8.4$ earthquake and 19 P and 17 SH waveforms for the $M_w = 7.9$ earthquake. The duration of the waveforms used for modelling the earthquake was 120 s for both the $M_w = 8.4$ earthquake and the $M_w = 7.9$ earthquake. Details on the GPS and InSAR data and the inversion method are given in Supplementary Information.

Received 30 May; accepted 17 October 2008.

1. Lay, T. *et al.* The great Sumatra-Andaman earthquake of 26 December 2004. *Science* **308**, 1127–1133 (2005).
2. Chlieh, M. *et al.* Coseismic slip and afterslip of the great M_w 9.15 Sumatra-Andaman earthquake of 2004. *Bull. Seismol. Soc. Am.* **97**, S152–S173 (2007).
3. Ammon, C. J. *et al.* Rupture process of the 2004 Sumatra-Andaman earthquake. *Science* **308**, 1133–1139 (2005).
4. Subarya, C. *et al.* Plate-boundary deformation associated with the great Sumatra-Andaman earthquake. *Nature* **440**, 46–51 (2006).
5. Briggs, R. W. *et al.* Deformation and slip along the Sunda Megathrust in the great 2005 Nias-Simeulue earthquake. *Science* **311**, 1897–1901 (2006).
6. Newcomb, K. & McCann, W. Seismic history and seismotectonics of the Sunda arc. *J. Geophys. Res.* **92**, 421–439 (1987).
7. Natawidjaja, D. H. *et al.* Source parameters of the great Sumatran megathrust earthquakes of 1797 and 1833 inferred from coral microatolls. *J. Geophys. Res.* **111**, doi:10.1029/2005JB004025 (2006).
8. Ruff, L. & Kanamori, H. Seismic coupling and uncoupling at subduction zones. *Tectonophysics* **99**, 99–117 (1983).
9. Pacheco, J. F., Sykes, L. R. & Scholz, C. H. Nature of seismic coupling along simple plate boundaries of the subduction type. *J. Geophys. Res.* **98**, 14133–14160 (1993).
10. Sieh, K., Ward, S. N., Natawidjaja, D. & Suwargadi, B. W. Crustal deformation at the Sumatran subduction zone revealed by coral rings. *Geophys. Res. Lett.* **26**, 3141–3144 (1999).
11. Burgmann, R. *et al.* Interseismic coupling and asperity distribution along the Kamchatka subduction zone. *J. Geophys. Res.* **110**, doi:10.1029/2005JB003648 (2005).
12. Suwa, Y., Miura, S., Hasegawa, A., Sato, T. & Tachibana, K. Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *J. Geophys. Res.* **111**, doi:10.1029/2004JB003203 (2006).
13. Freymueller, J. T. & Beavan, J. Absence of strain accumulation in the western Shumagin segment of the Alaska subduction zone. *Geophys. Res. Lett.* **26**, 3233–3236 (1999).

14. Pritchard, M. E. & Simons, M. An aseismic slip pulse in northern Chile and along-strike variations in seismogenic behavior. *J. Geophys. Res.* **111**, doi:10.1029/2006JB004258 (2006).
15. Chlieh, M., Avouac, J.-P., Sieh, K., Natawidjaja, D. H. & Galetzka, J. Heterogeneous coupling on the Sumatra megathrust constrained from geodetic and paleogeodetic measurements. *J. Geophys. Res.* **113**, doi:10.1029/2007JB004981 (2008).
16. Meltzner, A. J. *et al.* Coseismic, postseismic, and interseismic deformation, and long-term segmentation near the boundary of the 2004 and 2005 Sunda megathrust ruptures. *Eos* **88** (Fall meeting), abstr. S24A-02 (2007).
17. Natawidjaja, D. *et al.* Paleogeodetic records of seismic and aseismic subduction from central Sumatran microatolls, Indonesia. *J. Geophys. Res.* **109**, doi:10.1029/2003JB0002398 (2004).
18. Natawidjaja, D. H. *et al.* Interseismic deformation above the Sunda Megathrust recorded in coral microatolls of the Mentawai islands, West Sumatra. *J. Geophys. Res.* **112**, doi:10.1029/2006JB004450 (2007).
19. Thatcher, W. Order and diversity in the modes of Circum-Pacific earthquake recurrence. *J. Geophys. Res.* **95**, 2609–2623 (1990).
20. Ando, M. Source mechanisms and tectonic significance of historical earthquakes along Nankai Trough, Japan. *Tectonophysics* **27**, 119–140 (1975).
21. Lay, T., Kanamori, H. & Ruff, L. The asperity model and the nature of large subduction zone earthquakes. *Earthq. Predict. Res.* **1**, 3–71 (1982).
22. Rosen, P. A., Henley, S., Peltzer, G. & Simons, M. Updated repeat orbit interferometry package released. *Eos* **85**, 47 (2004).
23. Rundle, J. B. & Kanamori, H. Application of an inhomogeneous stress (patch) model to complex subduction zone earthquakes - a discrete interaction matrix approach. *J. Geophys. Res.* **92**, 2606–2616 (1987).
24. Perfettini, H. & Ampuero, J. P. Dynamics of a velocity strengthening fault region: Implications for slow earthquakes and postseismic slip. *J. Geophys. Res.* **113**, doi:10.1029/2007JB005398 (2008).
25. Liu, Y. & Rice, J. R. Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *J. Geophys. Res.* **110**, B08307 (2005).
26. Ben-Zion, Y. & Rice, J. R. Earthquake failure sequences along a cellular fault zone in a 3-dimensional elastic solid containing asperity and nonasperity regions. *J. Geophys. Res.* **98**, 14109–14131 (1993).
27. Cochard, A. & Madariaga, R. Complexity of seismicity due to highly rate-dependent friction. *J. Geophys. Res.* **101**, 25321–25336 (1996).
28. Park, S. C. & Mori, J. Are asperity patterns persistent? Implication from large earthquakes in Papua New Guinea. *J. Geophys. Res.* **112**, doi:10.1029/2006JB004481 (2007).
29. Ji, C., Wald, D. & Helmberger, D. V. Source description of the 1999 Hector Mine, California earthquake, Part I: Wavelet domain inversion theory and resolution analysis. *Bull. Seismol. Soc. Am.* **92**, 1192–1207 (2002).
30. Abercrombie, R. E., Antolik, M. & Ekstrom, G. The June 2000 M_w 7.9 earthquakes south of Sumatra: Deformation in the India-Australia Plate. *J. Geophys. Res.* **108**, doi:10.1029/2001JB000674 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements: This study was partly funded by the NSF (grant EAR-0538333) and the Gordon and Betty Moore Foundation. This is Caltech Tectonics Observatory contribution no. 93. We thank R. Burgmann for comments and suggestions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.-P.A. (avouac@gps.caltech.edu).

LETTERS

The pectoral fin of *Panderichthys* and the origin of digits

Catherine A. Boisvert¹, Elga Mark-Kurik² & Per E. Ahlberg¹

One of the identifying characteristics of tetrapods (limbed vertebrates) is the presence of fingers and toes. Whereas the proximal part of the tetrapod limb skeleton can easily be homologized with the paired fin skeletons of sarcopterygian (lobe-finned) fish, there has been much debate about the origin of digits. Early hypotheses¹ interpreted digits as derivatives of fin radials, but during the 1990s the idea gained acceptance that digits are evolutionary novelties without direct equivalents in fish fin skeletons. This was partly based on developmental genetic data², but also substantially on the pectoral fin skeleton of the elpistostegid (transitional fish/tetrapod) *Panderichthys*, which appeared to lack distal digit-like radials³. Here we present a CT scan study of an undisturbed pectoral fin of *Panderichthys* demonstrating that the plate-like 'ulnare' of previous reconstructions is an artefact and that distal radials are in fact present. This distal portion is more tetrapod-like than that found in *Tiktaalik*⁴ and, in combination with new data about fin development in basal actinopterygians⁵, sharks⁶ and lungfish⁷, makes a strong case for fingers not being a novelty of tetrapods but derived from pre-existing distal radials present in all sarcopterygian fish.

A near-complete specimen of *Panderichthys* from the late Middle Devonian period (385 million years ago; see Supplementary Information) of Lode, Latvia (Institute of Geology at Tallinn University of Technology specimen number GIT434-1) forms the basis of this study. Although the dorsal part of the skull and left side of the body have suffered substantial damage from a mechanical excavator, the specimen was originally well preserved. Notably, its body axis is straight (determined from the alignment of dorsal mid-line scales and the symmetry plane of the skull) and it appears less dorsoventrally compressed than others from the same locality. This is shown, for example, by the narrow skull outline and near-vertical cheek-plate fragments, distinctly different from the flattened and splayed skulls that have been published^{8,9}. The specimen also contains the only known pelvis and pelvic fin skeleton of *Panderichthys*¹⁰.

The right pectoral fin and most of the shoulder girdle are preserved in articulation, with the fin concealed under the body. This region of the specimen was CT scanned at the East-Tallinn Central Hospital (see Methods Summary and Supplementary Methods). The scanned region comprises the entire fin endoskeleton and an estimated 40% of the lepidotrichial fin web, as well as most of the shoulder girdle. The presence of X-ray-reflective crystal growths in the shoulder region unfortunately prevented complete modelling of the scapulocoracoid. The entire fin is covered by scales and lepidotrichia, which cannot be modelled individually but are easily separated from the darker endoskeleton. The leading or preaxial margin of the fin is dipping ventrolaterally into the substrate (Fig. 1a–c). In contrast to early tetrapods, in which the limb projects at an angle from the body wall¹¹, the fin of *Panderichthys* is oriented anteroposteriorly in line

with the body axis. In distal cross section (Fig. 1c), the fin is slightly cambered, convex dorsally and concave ventrally. The dorsal part of the cleithrum is missing but the notably long anocleithrum is complete (Fig. 1b, d). The proximal quarter of the anocleithrum is overlapped by the supracleithrum.

The fin endoskeleton is complete and is composed of the humerus, radius, ulna, ulnare, bipartite intermedium and four small distal radials arranged in a transverse terminal array (Fig. 2). The humerus closely resembles the most recent published description¹² as far as the morphological details are concerned, but it is much less flattened, displaying a pear-shaped proximal cross section rather than the

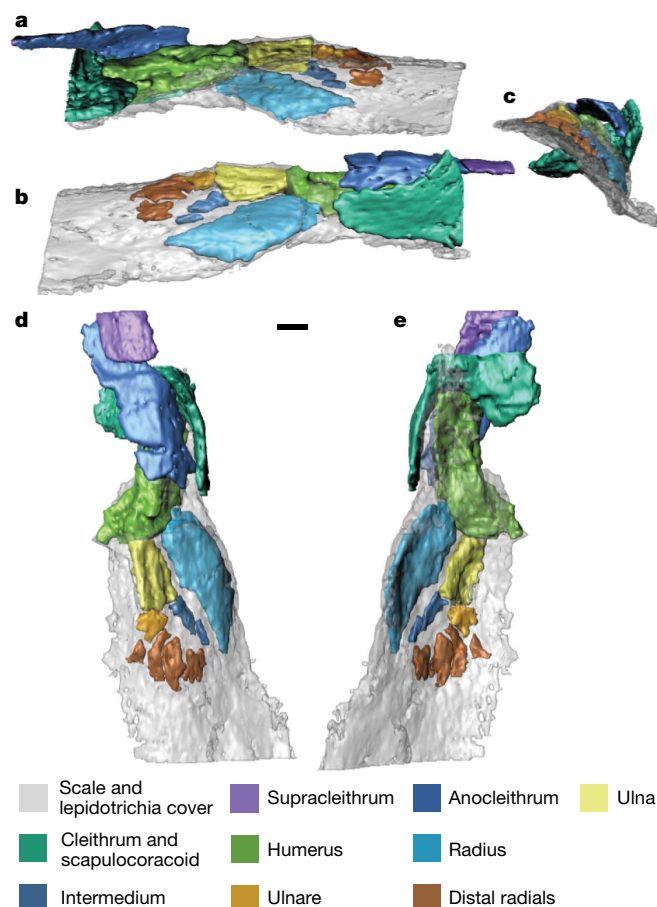


Figure 1 | Pectoral fin, girdle and scale plus lepidotrichia cover of *Panderichthys*. **a**, Postaxial view; **b**, preaxial view; **c**, distal cross section; **d**, dorsal view; and **e**, ventral view. Scale bar, 1 cm.

¹Subdepartment of Evolutionary Organismal Biology, Department of Physiology and Developmental Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18A, 752 36 Uppsala, Sweden. ²Institute of Geology at Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia.

reported very dorsoventrally flattened shape¹³. We attribute this difference to the unusually three-dimensional preservation of GIT434-1. In most respects, the humerus agrees well with that of *Tiktaalik*⁴. However, the orientation of the scapulohumeral joint is

parallel to the scan plane, and the thickness of scan slices (0.6 mm) prevents us from reconstructing the joint surfaces in detail.

The ectepicondyle is linear, low and parallel to the post-axial edge of the humerus (Fig. 2a). It is shorter than in *Tiktaalik*⁴, ending some distance from the ulnar articulation. On the ventral surface, the humeral ridge is slightly oblique (Fig. 2b) and much higher than previously described¹², terminating at the base of the entepicondyle rather than extending onto its ventral face as in *Tiktaalik*^{4,13}. The radius is a sickle-shaped blade, very slender in axial aspect and convex ventrally. It is smooth on both the dorsal and the ventral surfaces and has a thickened lateral edge. Like in *Tiktaalik*⁴, the ulna has a complex shape with several longitudinal grooves and ridges.

The most distal portion of the fin endoskeleton had been described as being composed of only two elements—a large plate-like ulnare and a narrow intermedium in close contact³—but examination by CAB of the published material (Palaeontological Institute of the Russian Academy of Sciences, Moscow specimen numbers PIN 3547/26 and PIN 3547/18) showed that the region is covered by a thin sediment film bearing impressions of the inner faces of scales, which conceals any endoskeletal elements underneath. These scale impressions can in fact be discerned in the published reconstruction drawings as a regular pattern on the ‘ulnare’³. Our CT scan thus reveals the distal fin endoskeleton of *Panderichthys* for the first time. The ulnare is small and flattened, articulating proximally with the ulna and distally with two terminal radials. The intermedium is a flattened rod, narrower than previously described, articulating directly in line to the lateral ridge of the ulna. It appears to consist of separate proximal and distal components as in *Tiktaalik*⁴, but the resolution of the ‘joint’ is poor and it is possible that this divide is an artefact.

The distal-most part of the fin is composed of four distal radials, arranged in a slight arc distal to the ulnare and the intermedium. The separation of radials from each other and from the ulnare can be seen by density differences in the scan (see Supplementary Information) but also through slight differences in the orientation of the elements. All radials have flattened cross-sections of the same thickness but they vary in width. If the axis of the fin is interpreted to extend through the middle of the ulnare, one of the distal radials is postaxial, one is in line with the ulnare and two are preaxial. The two most preaxial distal radials do not articulate directly with an axial element (Fig. 2).

Our reinterpretation of the distal fin endoskeleton of *Panderichthys* removes the final piece of evidence supporting the formerly popular hypothesis that tetrapod digits are wholly new structures without homologues in sarcopterygian fish fins^{2,14–16}. This hypothesis, which was based partly on the complete absence of plausible digit homologues in *Panderichthys* (then the closest known relative of tetrapods)³, has already been called into question by the discovery of digit-like radials in *Tiktaalik*⁴ and the fact that *Hox* gene expression patterns closely resembling those associated with digit formation in tetrapods occur in the distal fin skeletons of paddlefish⁵ and Australian lungfish⁷. Our new data show that *Panderichthys* is not an anomaly: like *Tiktaalik* and other fish members of the Tetrapodomorpha^{17,18} it has distal radials that can be interpreted as digit homologues.

The pectoral fin skeletons of *Panderichthys* and *Tiktaalik* share certain unusual features such as a blade-like radius and a longitudinal ridge-and-groove on the flexor surface of the ulna. These can tentatively be interpreted as attributes of the ‘elpistostegid’ segment of the tetrapod stem lineage and thus ancestral for the tetrapod forelimb. Given that recent phylogenies consistently place *Panderichthys* below *Tiktaalik* in the tetrapod stem group^{19,20}, it is surprising to discover that its pectoral fin skeleton is more limb-like than that of its supposedly more derived relative. In *Tiktaalik*⁴, like in ‘osteolepiforms’¹⁷ and rhizodonts¹⁸ (more primitive fish members of the stem group), the ulna and ulnare are of similar size. The axis of the fin comprises two more elements distal to the ulnare, and the distal radials are arranged pinnately around this axis. In contrast, in *Panderichthys* and tetrapods, the ulna is much longer than the ulnare, the ulnare

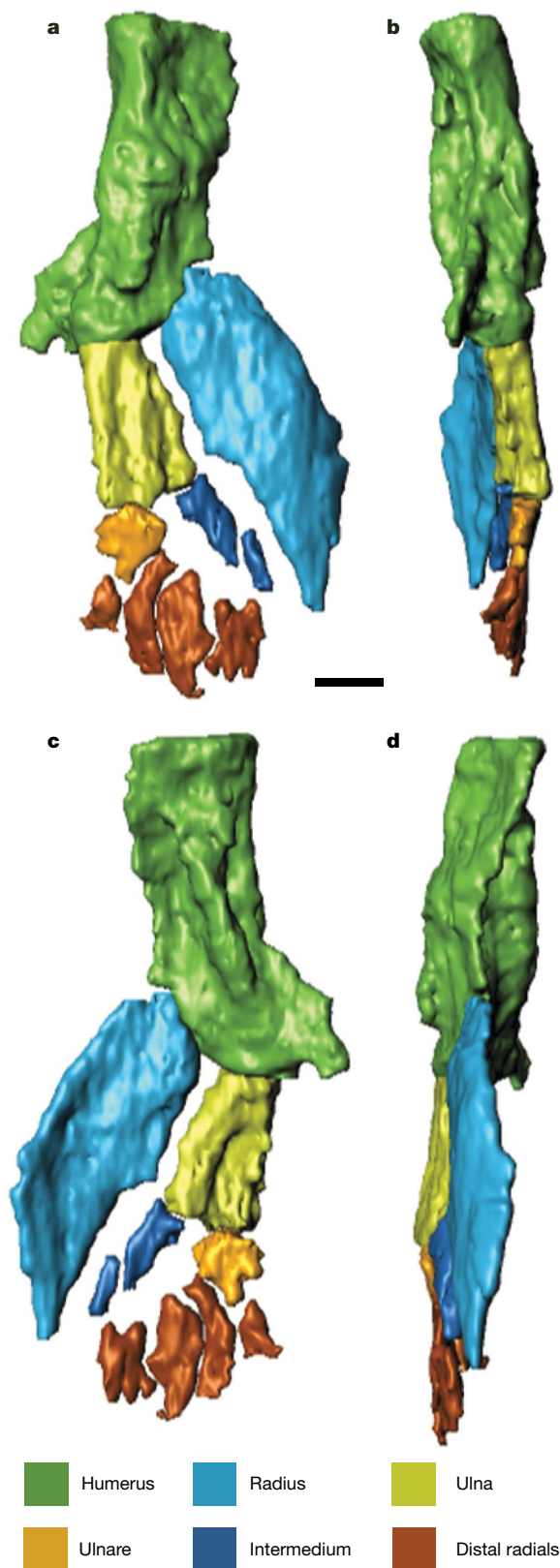


Figure 2 | Endoskeleton of the pectoral fin of *Panderichthys rhombolepis*. a, Dorsal; b, postaxial; c, ventral; and d, preaxial views. Scale bar, 1 cm.

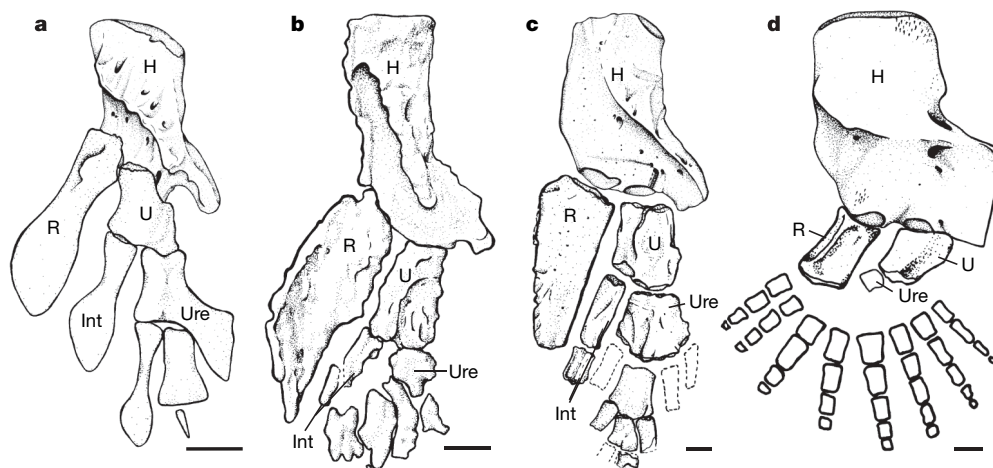


Figure 3 | Pectoral endoskeleton of tetrapodomorph fish and a tetrapod in ventral view. Pectoral fins of: **a**, *Eusthenopteron*; **b**, *Panderichthys*; and **c**, *Tiktaalik*. **d**, Limb of *Acanthostega*. H, humerus; Int, intermedium;

R, radius; U, ulna; Ure, ulnare. Scale bar, 1 cm. **a** is redrawn from ref. 17, **c** from ref. 4 and **d** from ref. 11.

is the last axial element, and the distal radials/digits are arranged in a transverse fan shape^{11,21} (Fig. 3). It is difficult to say whether this character distribution implies that *Tiktaalik* is autapomorphic, that *Panderichthys* and tetrapods are convergent, or that *Panderichthys* is closer to tetrapods than *Tiktaalik*. At any rate, it demonstrates that the fish–tetrapod transition was accompanied by significant character incongruence in functionally important structures.

METHODS SUMMARY

The specimen was CT scanned on a medical scanner at East-Tallinn Central Hospital, Estonia, and modelled independently using Mimics v.9.11 (Materialise NV, <http://www.materialise.com/mimics>) and rendered using Rhinoceros (version 3.0 SR4) with the plug-in Bongo (version 1.0).

Received 12 June; accepted 14 August 2008.

Published online 21 September 2008.

1. Gegenbaur, C. *Grundriss der vergleichenden anatomie* (Wilhelm Engelmann, 1874).
2. Sordino, P., van der Hoeven, F. & Duboule, D. *Hox* gene expression in teleost fins and the origin of vertebrate digits. *Nature* **375**, 678–681 (1995).
3. Vorobyeva, E. I. The role of development and function in formation of tetrapod-like pectoral fins. *Zh. Obshch. Biol.* **53**, 149–158 (1992).
4. Shubin, N. H., Daeschler, E. B. & Jenkins, F. A. Jr. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* **440**, 764–771 (2006).
5. Davis, M. C., Dahn, R. D. & Shubin, N. H. An autopodial-like pattern of *Hox* expression in the fins of a basal actinopterygian fish. *Nature* **447**, 473–476 (2007).
6. Dahn, R. D. *et al.* Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning. *Nature* **445**, 311–314 (2007).
7. Johanson, Z. *et al.* Fish fingers: digit homologues in sarcopterygians fish fins. *J. Exp. Zool. (Mol. Dev. Evol.)* **308B**, 757–768 (2007).
8. Vorobyeva, E. I. Observations on two rhipidistian fishes from the Upper Devonian of Lode, Latvia. *Zool. J. Linn. Soc.* **70**, 191–201 (1980).
9. Vorobyeva, E. I. & Schultze, H.-P. In *Origins of the Higher Groups of Tetrapods: Controversy and Consensus* (eds Schultze, H.-P. & Trueb L.) 68–109 (Cornell Univ. Press, 1991).
10. Boisvert, C. A. The pelvic fin and girdle of *Panderichthys* and the origin of tetrapod locomotion. *Nature* **438**, 1145–1147 (2005).

11. Coates, M. I. The Devonian tetrapod *Acanthostega gunnari* Jarvik: postcranial anatomy, basal interrelationships and patterns of skeletal evolution. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
12. Vorobyeva, E. I. Morphology of the humerus in the rhipidistian *Crossopterygii* and the origin of tetrapods. *Paleontol. J.* **34**, 632–641 (2000).
13. Boisvert, C. A. in *Forty Years of Early Vertebrates: Papers from the 11th International Symposium on Early and Lower Vertebrates* (eds Ahlberg, P. E., Blom H. & Boisvert, C. A.) (Vol. Acta Zool. special volume, Blackwell, in the press).
14. Shubin, N. H., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
15. Gregory, W. K. & Raven, H. C. Studies on the origin and early evolution of paired fins and limbs. *Ann. NY Acad. Sci.* **42**, 273–360 (1941).
16. Westoll, T. S. The origin of the primitive tetrapod limb. *Proc. R. Soc. Lond. B* **131**, 373–393 (1943).
17. Andrews, M. S. & Westoll, T. S. The postcranial skeleton of *Eusthenopteron foordi* Whiteaves. *Trans. R. Soc. Edinb.* **68**, 207–328 (1970).
18. Davis, M. C., Shubin, N. H. & Daeschler, E. B. A new specimen of *Sauripterus taylori* (Sarcopterygii, Osteichthyes) from the Famennian Catskill Formation of North America. *J. Vert. Paleontol.* **24**, 26–40 (2004).
19. Daeschler, E. B., Shubin, N. H. & Jenkins, F. A. Jr. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature* **440**, 757–763 (2006).
20. Long, J. A. *et al.* An exceptional Devonian fish from Australia sheds light on tetrapod origins. *Nature* **444**, 199–202 (2006).
21. Jarvik, E. The Devonian tetrapod *Ichthyostega*. *Fossils and Strata* **40**, 1–213 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Ross and the team at East-Tallinn Central Hospital, Estonia for the CT-scanning, E. Vorobyeva for permission to examine the *Panderichthys* material in her care, N. Shubin for access to *Tiktaalik*, and M. I. Coates for valuable discussions. C.A.B. and P.E.A. are supported by Vetenskapsrådet.

Author Contributions E.M.-K wrote one of the Supplementary Information files and arranged the logistics for the CT scan, P.E.A. and E.M.-K. scanned the specimen, and C.A.B. and P.E.A. produced the models and descriptions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.A.B. (catherine.boisvert@ebc.uu.se).

Direct control of paralysed muscles by cortical neurons

Chet T. Moritz¹, Steve I. Perlmuter¹ & Eberhard E. Fetz¹

A potential treatment for paralysis resulting from spinal cord injury is to route control signals from the brain around the injury by artificial connections. Such signals could then control electrical stimulation of muscles, thereby restoring volitional movement to paralysed limbs^{1–3}. In previously separate experiments, activity of motor cortex neurons related to actual or imagined movements has been used to control computer cursors and robotic arms^{4–10}, and paralysed muscles have been activated by functional electrical stimulation^{11–13}. Here we show that *Macaca nemestrina* monkeys can directly control stimulation of muscles using the activity of neurons in the motor cortex, thereby restoring goal-directed movements to a transiently paralysed arm. Moreover, neurons could control functional stimulation equally well regardless of any previous association to movement, a finding that considerably expands the source of control signals for brain-machine interfaces. Monkeys learned to use these artificial connections from cortical cells to muscles to generate bidirectional wrist torques, and controlled multiple neuron–muscle pairs simultaneously. Such direct transforms from cortical activity to muscle stimulation could be implemented by autonomous electronic circuitry, creating a relatively natural neuroprosthesis. These results are the first demonstration that direct artificial connections between cortical cells and muscles can compensate for interrupted physiological pathways and restore volitional control of movement to paralysed limbs.

Spinal cord injury impairs neural pathways between the brain and limbs, but spares both the motor cortex and muscles. Recent studies have shown that quadriplegic patients could volitionally modulate activity of neurons in the hand area of the motor cortex, even after several years of paralysis⁶, and that monkeys could use cortical activity to control a robotic arm to acquire targets⁴ and feed themselves⁵. These and other brain–machine interface studies used sophisticated algorithms to decode task-related activity of neural populations and to calculate requisite control parameters for external devices^{4–6,8–10}. An alternative strategy to restore limb function is to directly connect cortical cell activity to control the stimulation of a patient's paralysed muscles (Fig. 1a). Here we show that monkeys can learn to use direct artificial connections from arbitrary motor cortex cells to grade stimulation delivered to multiple muscles and restore goal-directed movement to a paralysed arm.

In previous biofeedback studies monkeys rapidly learned to control the discharge rates of newly isolated neurons in the motor cortex to obtain rewards^{14,15}. We used similar operant conditioning techniques for single neurons in the hand and wrist area of the motor cortex of two monkeys (see Methods and Supplementary Information). We tested volitional control of cell activity by displaying smoothed discharge rate as cursor position on a monitor and rewarding the monkeys for maintaining activity within randomly presented high- or low-rate targets. The directional tuning of most cells was also characterized in an isometric two-dimensional wrist

target-tracking task. Our experiment, however, used all sufficiently well-isolated cells encountered, with no selection bias for possible association to movement or directional tuning.

Monkeys demonstrated volitional control of the discharge rates of nearly all cells tested within the first 10-min practice session. Although cell activity controlled the cursor directly, monkeys often continued to produce wrist torques during these initial sessions (Supplementary Fig. 1). We then blocked peripheral nerves innervating the wrist muscles with a local anaesthetic (see Methods). Despite loss of motor function and sensory feedback from the innervated forearm, monkeys continued to control the cursor with cell

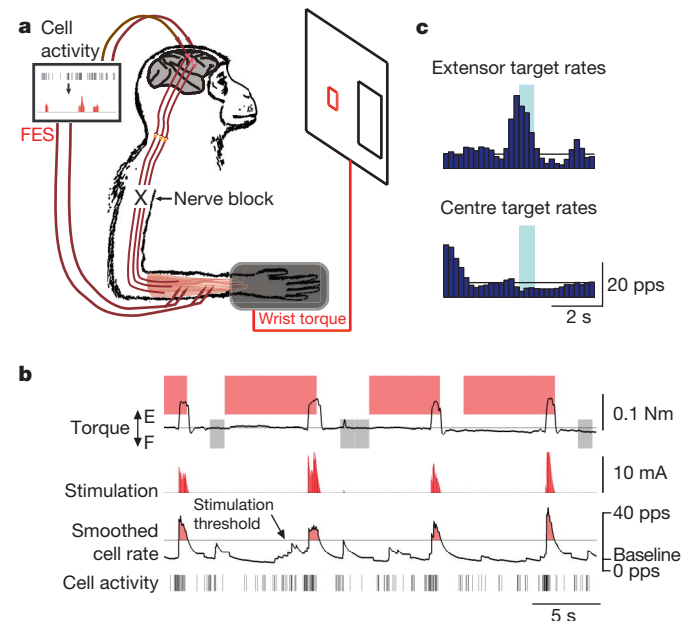


Figure 1 | Brain-controlled functional electrical stimulation (FES) of muscle. **a**, Schematic shows cortical cell activity converted to FES during peripheral nerve block. **b**, Example of motor cortex cell activity controlling FES of paralysed wrist extensors. Extensor (red shading) and centre (grey shading) wrist torque targets were randomly presented. Monkeys learned to modulate smoothed cell rate to control proportional muscle stimulation. FES was delivered to muscles EDC and ED4,5 at 50 s^{-1} , with current proportional to cell rate above a stimulation threshold ($0.4 \text{ mA pps}^{-1} \times [\text{cell rate} - 16 \text{ pps}]$; $\leq 10 \text{ mA}$). Here pps indicates pulses per second. **c**, Histograms of cell rates while acquiring the extensor and centre targets, illustrating cell activity used to successfully control FES. Shading indicates target hold period and horizontal line denotes baseline cell rate.

¹Department of Physiology & Biophysics and Washington National Primate Research Center, University of Washington, Seattle, Washington 98195, USA.

activity for 45 out of 46 cells after the nerve block. Supplementary Fig. 1 shows the loss of flexor and extensor torques after injections of local anaesthetic, while the monkey continued to volitionally control the cell activity. The nerve block was confirmed by the inability of the monkeys to perform the two-dimensional torque-tracking task.

We then converted cell activity into proportional stimuli delivered to paralysed muscles, generating functional electrical stimulation (FES). The cursor was now controlled by wrist torque, and the monkeys were rewarded for maintaining FES-evoked torque within peripheral and centre (that is, zero-torque) targets for 0.5–1.0 s. To allow the monkeys to grade contraction force, stimulation current was made linearly proportional to cell rate when the cell discharged above a threshold.

The example in Fig. 1b, c shows a monkey modulating cell activity to generate appropriate torques by controlling FES of paralysed wrist extensor muscles. The monkey learned to increase cell activity to activate the stimulator and acquire the extensor targets, and to maintain activity below the stimulation threshold to relax the muscle and acquire the centre targets. Both monkeys were able to control muscle FES during nerve block and acquire torque targets with 44 out of the 45 cells tested (5 cells from monkey I and 39 from monkey L).

For each cell the monkeys' control improved with practice, as evidenced by more rapid acquisition of targets and fewer errors. Monkeys began using cell activity to control the stimulator almost immediately, and improved substantially during the relatively brief practice sessions with each cell (mean duration 66 min). To quantify this improvement we compared performance during the initial two minutes of practice and during the two-minute period with the highest performance, typically just before task difficulty was increased to probe the limits of FES control. The rate of target acquisition with FES control was over three times greater during peak performance (14.1 ± 5.3 torque targets acquired per min; mean \pm s.d.) compared to the beginning of practice (4.0 ± 4.3 targets per min; $P < 0.001$; Supplementary Fig. 2). Peak target acquisition rates during brain-controlled FES were similar to those seen when cell activity controlled the cursor directly before nerve block (13.2 ± 5.5 targets per min; $P = 0.66$).

With continued practice monkeys also learned to control the torque more precisely with cell activity, making fewer target acquisition errors and often acquiring targets on the first attempt. A target acquisition error was defined as triggering the stimulator to acquire the peripheral target when the centre target was displayed. Monkeys made target errors on only $0.8 \pm 5.1\%$ of targets during peak performance for each cell compared to $20.7 \pm 28.9\%$ of targets at the beginning of practice ($P < 0.001$; Supplementary Fig. 3). They also made 81% fewer failed attempts to acquire the target during peak performance (0.10 ± 0.31 failed attempts per target) compared to the beginning of practice (0.52 ± 0.93 ; $P < 0.001$).

To test whether FES could also be controlled by decreases in cell activity, we set stimulation current to be inversely proportional to cell rate below a threshold for 11 cells. Monkey L learned to control stimulation with this inverse relation just as well as with a positive relation between cell rate and stimulus current (38 cells, some tested in both groups; $P > 0.46$), acquiring 13.4 ± 3.9 targets per minute and making no errors during peak performance.

The activity of a single cell could also be used to control stimulation of antagonist muscle groups and restore bidirectional movements. Figure 2 shows an example of one cell that controlled stimulation of flexor muscles with high discharge rates and extensor muscles with low rates. The monkey learned to control cell activity and grade contraction force to rapidly satisfy targets at five different torque levels. The nerve blocks remained very effective, as evidenced by negligible torques produced in either direction when the stimulators were turned off during target presentation (Fig. 2b). Seven cells tested with such bidirectional control performed similarly to cells that controlled only one muscle group, although target acquisition rates were marginally slower (9.8 ± 3.7 targets per min; $P = 0.06$).

The assumptions underlying common neural decoding schemes would predict that monkeys should be able to control FES torque better with cells that are strongly related to wrist movements than with unrelated cells. To investigate this, we documented cell activity during a two-dimensional wrist target-tracking task before the nerve

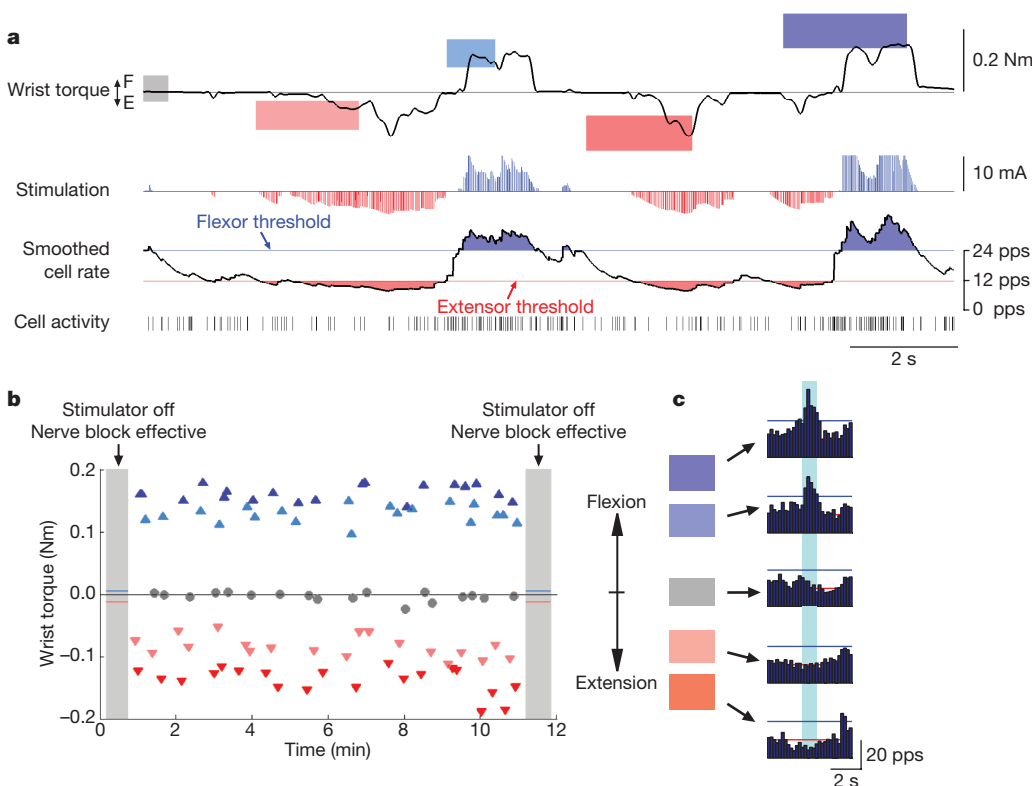
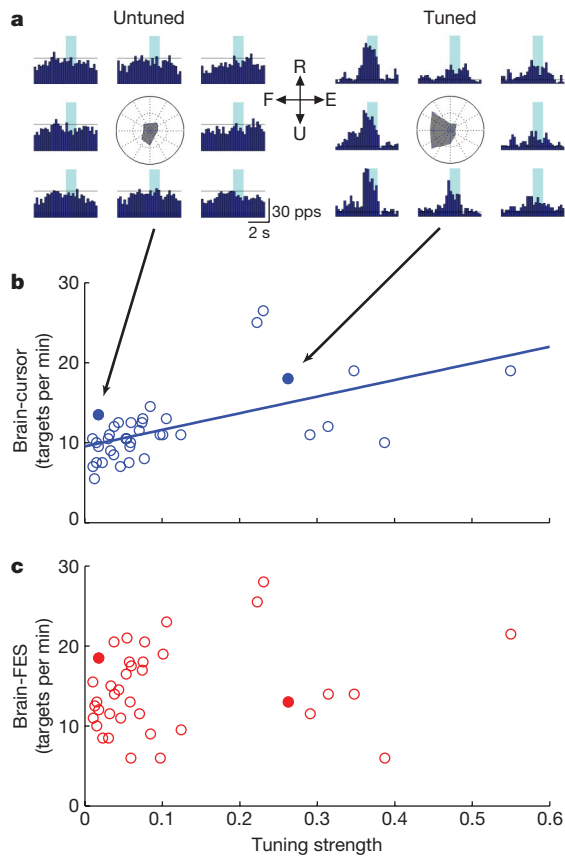


Figure 2 | Brain-controlled FES of multiple muscles restores graded torque in two directions. **a**, The monkey acquired targets at five levels of flexion-extension (F-E) torque using the activity of a single cell to grade FES delivered to both flexor (FCU) and extensor (ECU and ED4,5) muscles. Flexor FES was proportional to the rate above a threshold ($0.8 \text{ mA pps}^{-1} \times [\text{cell rate} - 24 \text{ pps}]$; $\leq 10 \text{ mA}$); extensor FES was inversely proportional to the cell rate below a second threshold ($0.6 \text{ mA pps}^{-1} \times [12 \text{ pps} - \text{cell rate}]$; $\leq 10 \text{ mA}$). **b**, Average torques produced to satisfy the five targets during 12 min of practice. With the stimulator off (shaded periods) the monkey could not produce torques greater than 10% of magnitudes used to acquire the targets (blue and red lines), confirming the efficacy of nerve block. **c**, Histograms of cell rate used to acquire five target levels (coloured boxes at left). Horizontal lines indicate FES thresholds for flexor (blue) and extensor (red) stimulation.



block, and calculated the directional tuning for each cell (Fig. 3a). The magnitude of directional tuning did correlate significantly with the ability of the monkeys to bring the cursor into the optimally placed targets with cell activity during the initial 10-min practice period ($r^2 = 0.33$, $P < 0.001$; Fig. 3b). However, cell tuning was not a good predictor of the peak target acquisition rates during

Figure 3 | Cell directional tuning is unrelated to FES control. **a**, Responses of an untuned and a strongly tuned cell (solid symbols in **b** and **c**). The surrounding peri-event histograms show cell activity while matching (shading) each of eight peripheral torque targets in the flexion-extension (F-E) and radial-ulnar (R-U) plane during the unparalysed tracking task (horizontal lines denote baseline cell rates). The radial plot at the centre summarizes cell activity while matching each peripheral target. **b**, **c**, Maximum target acquisition rates during direct brain control of cursor (**b**) and brain-controlled FES (**c**) plotted as a function of directional tuning strength for cells recorded during the torque-tracking task ($n = 38$). Performance controlling a cursor directly with cell activity was significantly correlated with cell tuning (**b**; $r^2 = 0.33$, $P < 0.001$). Subsequent brain-controlled FES performance was uncorrelated with cell tuning (**c**; $r^2 = 0.03$, $P = 0.33$).

subsequent brain-controlled FES ($r^2 = 0.03$, $P = 0.33$; Fig. 3c). For example, with the untuned cell on the left in Fig. 3a the monkey acquired 18.5 targets per minute. The tuned ($n = 9$) and untuned ($n = 29$) cells showed no differences in any measure of FES control (target acquisition rates, errors or failed attempts; $P > 0.51$).

Extending the strategy of direct neural control to more complex movements will require further control signals. As a first step towards this goal, we tested a monkey's ability to simultaneously control two cell-muscle pairs. Figure 4 shows monkey L using high discharge rates of one cell to control FES of flexor muscles and high rates of a second cell to control extensor muscles. The monkey learned to independently modulate the activity of five cell pairs in order to control antagonist muscles and rapidly acquire bidirectional torque targets at rates similar to single cells (11.6 ± 3.8 targets per min, $P = 0.32$).

These findings have several implications for future approaches to neuroprosthetic control. In contrast to the conventional strategy of deriving control signals from the combined activity of a neural population^{4-6,8-10}, it may prove efficacious to maintain separate signal pathways from cells to muscles. Using direct channels from single cells to specific muscles may provide the brain with more distinguishable outcomes of the cell activity¹⁶ and allow innate motor learning mechanisms to help optimize control of the new connections. The ability of the brain to adapt to new but consistent sensorimotor contingencies has been amply documented^{17,18}, and motor cortex can adapt rapidly to learn new motor skills^{19,20}. Motor circuitry can compensate for drastic changes in connectivity, such as surgically

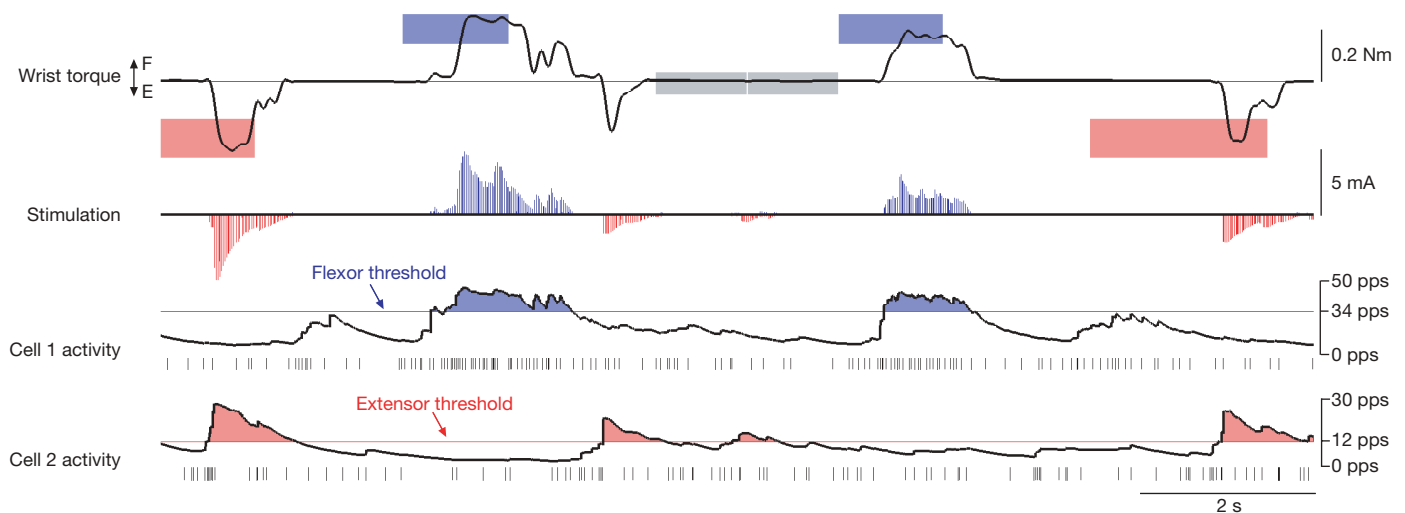


Figure 4 | Two neurons control FES. Monkey L simultaneously modulated activity of two neurons, each controlling proportional stimulation of a different muscle group when above a threshold. Monkey L acquired randomly presented flexor (blue), extensor (red) and centre (grey) targets by

using cell 1 to stimulate a flexor muscle (FCU; $0.2 \text{ mA pps}^{-1} \times [\text{cell rate} - 34 \text{ pps}]$) and cell 2 to stimulate extensor muscles (ECU and ED4,5; $0.4 \text{ mA pps}^{-1} \times [\text{cell rate} - 12 \text{ pps}]$).

cross-connected nerves controlling wrist flexor and extensor muscles²¹, or targeted reinnervation for control of prosthetic limbs²².

Our finding that monkeys could learn to use virtually any motor cortex cell to control muscle stimulation—regardless of the cell's original relation to wrist movement (Fig. 3c)—suggests another advantage of directly tapping single cell activity. Strategies based on decoding the activity of neural ensembles to obtain movement parameters or muscle activity depend on finding cells that modulate sufficiently with the output variables during actual or imagined movements^{4–6,8–10}. Instead, arbitrary cells available on recording arrays could be brought under volitional control using biofeedback, substantially expanding the source of control signals for brain-machine interfaces. This and previous biofeedback studies^{14,15} have shown that even cells with no discernable relation to muscles can be volitionally modulated after brief practice sessions. Issues concerning the use of individual cells and neural populations for prosthetic control are further discussed in Supplementary Information.

The degree of FES control demonstrated here was limited by the relatively brief training time provided by the transient nerve block. Implanted electronic circuitry will enable adaptive learning over much longer times and under more varied conditions¹. For example, the autonomous 'Neurochip' system can discriminate single cell activity and deliver stimulation through days of free behaviour^{23,24}. In several preliminary FES sessions, we confirmed that this system would allow a monkey to trigger stimulation of a paralysed muscle with cell activity and acquire torque targets (Supplementary Fig. 4). Such autonomous low-power circuits could permit subjects to practice continuously with an artificial connection from brain to muscles or the spinal cord^{1,25–27}, without requiring complex decoding algorithms or robotic arms. Further development of such direct-control strategies may lead to implantable devices that could help restore volitional movements to individuals living with paralysis.

METHODS SUMMARY

Subjects. Two male *Macaca nemestrina* monkeys participated in the experiments (4–5-yr-old, weight 4.5–6.5 kg). All procedures were approved by the University of Washington Institutional Animal Care and Use Committee.

Recording and paralysis. The activity of single motor cortex cells was recorded using either acute (Monkey I and L) or chronic (Monkey L) electrodes. Each session began by quantifying the response of cells during an isometric, eight-target wrist torque-tracking task. Volitional control of cell activity was confirmed by operantly rewarding acquisition of targets with a cursor controlled by cell rates. Wrist muscles were then paralysed by injecting anaesthetic (3% chloroprocaine or 2% lidocaine, each with 1:100,000 adrenaline) into catheters or cuffs surrounding the median, ulnar and/or radial nerves.

Brain-controlled FES. Cell activity controlled the intensity of stimuli delivered by means of bipolar electrodes implanted in one or more paralysed wrist muscles. When cell rate (smoothed over a 0.5-s sliding window) crossed a threshold, biphasic constant-current stimuli (cathode leading; 0.75–1.0 ms pulse width) were delivered at 50 s⁻¹. For most cells, stimulus current was made proportional to cell rate above a threshold to allow the monkey to grade contraction force (for example, stimulus current = 0.1 mA × [cell rate – threshold]; to a maximum of 10 mA). Some cells controlled stimulation in inverse proportion to cell rate below a threshold.

Analysis. The strength of directional tuning was calculated for cells during the initial torque-tracking task using the vector method²⁸ (see Supplementary Information). Peak performance was quantified by the maximum number of targets acquired during a two-min period. Peak performance was compared among conditions and to performance during the initial two minutes of practice using the non-parametric rank sum test. Regression analysis determined correlations between directional tuning and peak performance during brain control of a cursor or FES.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 July; accepted 11 September 2008.

Published online 15 October 2008.

1. Jackson, A., Moritz, C. T., Mavoori, J., Lucas, T. H. & Fetz, E. E. The Neurochip BCI: towards a neural prosthesis for upper limb function. *IEEE Trans. Neural Syst. Rehabil. Eng.* **14**, 187–190 (2006).

2. Lauer, R. T., Peckham, P. H. & Kilgore, K. L. EEG-based control of a hand grasp neuroprosthesis. *Neuroreport* **10**, 1767–1771 (1999).
3. Fagg, A. H. et al. Biomimetic brain machine interfaces for the control of movement. *J. Neurosci.* **27**, 11842–11846 (2007).
4. Carmena, J. M. et al. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol.* **1**, E42 (2003).
5. Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S. & Schwartz, A. B. Cortical control of a prosthetic arm for self-feeding. *Nature* **453**, 109–1101 (2008).
6. Hochberg, L. R. et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* **442**, 164–171 (2006).
7. Kennedy, P. R., Bakay, R. A., Moore, M. M., Adams, K. & Goldwain, J. Direct control of a computer from the human central nervous system. *IEEE Trans. Rehabil. Eng.* **8**, 198–202 (2000).
8. Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* **305**, 258–262 (2004).
9. Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R. & Donoghue, J. P. Instant neural control of a movement signal. *Nature* **416**, 141–142 (2002).
10. Taylor, D. M., Tillery, S. I. & Schwartz, A. B. Direct cortical control of 3D neuroprosthetic devices. *Science* **296**, 1829–1832 (2002).
11. Nannini, N. & Horsch, K. Muscle recruitment with intrafascicular electrodes. *IEEE Trans. Biomed. Eng.* **38**, 769–776 (1991).
12. Peckham, P. H. et al. An advanced neuroprosthesis for restoration of hand and upper arm control using an implantable controller. *J. Hand Surg.* **27**, 265–276 (2002).
13. Stein, R. B., Aoyagi, Y., Mushahwar, V. K. & Prochazka, A. Limb movements generated by stimulating muscle, nerve and spinal cord. *Arch. Ital. Biol.* **140**, 273–281 (2002).
14. Fetz, E. E. & Baker, M. A. Operantly conditioned patterns on precentral unit activity and correlated responses in adjacent cells and contralateral muscles. *J. Neurophysiol.* **36**, 179–204 (1973).
15. Fetz, E. E. & Finocchio, D. V. Correlations between activity of motor cortex cells and arm muscles during operantly conditioned response patterns. *Exp. Brain Res.* **23**, 217–240 (1975).
16. Fetz, E. E. Volitional control of neural activity: implications for brain-computer interfaces. *J. Physiol. (Lond.)* **579**, 571–579 (2007).
17. Shadmehr, R. & Mussa-Ivaldi, F. A. Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **14**, 3208–3224 (1994).
18. Thach, W. T. Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. *J. Neurophysiol.* **41**, 654–676 (1978).
19. Gandolfo, F., Li, C., Benda, B. J., Schioppa, C. P. & Bizzi, E. Cortical correlates of learning in monkeys adapting to a new dynamical environment. *Proc. Natl Acad. Sci. USA* **97**, 2259–2263 (2000).
20. Nudo, R. J., Milliken, G. W., Jenkins, W. M. & Merzenich, M. M. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* **16**, 785–807 (1996).
21. Brinkman, C., Porter, R. & Norman, J. Plasticity of motor behavior in monkeys with crossed forelimb nerves. *Science* **220**, 438–440 (1983).
22. Kuiken, T. A., Dumanian, G. A., Lipschutz, R. D., Miller, L. A. & Stubblefield, K. A. The use of targeted muscle reinnervation for improved myoelectric prosthesis control in a bilateral shoulder disarticulation amputee. *Prosthet. Orthot. Int.* **28**, 245–253 (2004).
23. Jackson, A., Mavoori, J. & Fetz, E. E. Long-term motor cortex plasticity induced by an electronic neural implant. *Nature* **444**, 56–60 (2006).
24. Mavoori, J., Jackson, A., Diorio, C. & Fetz, E. An autonomous implantable computer for neural recording and stimulation in unrestrained primates. *J. Neurosci. Methods* **148**, 71–77 (2005).
25. Lemay, M. A. & Grill, W. M. Modularity of motor output evoked by intraspinal microstimulation in cats. *J. Neurophysiol.* **91**, 502–514 (2004).
26. Mushahwar, V. K., Gillard, D. M., Gauthier, M. J. & Prochazka, A. Intraspinal micro stimulation generates locomotor-like and feedback-controlled movements. *IEEE Trans. Neural Syst. Rehabil. Eng.* **10**, 68–81 (2002).
27. Tresch, M. C. & Bizzi, E. Responses to spinal microstimulation in the chronically spinalized rat and their relationship to spinal systems activated by lowthreshold cutaneous stimulation. *Exp. Brain Res.* **129**, 401–416 (1999).
28. Batschelet, E. *Circular Statistics in Biology* 3–39 (Academic, 1981).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Shupe for programming assistance, C. Kent and L. Miller for advice on nerve block, C. Kirby, A. Price and K. McElwain for animal care, and A. Jackson, Y. Nishimura and A. Richardson for comments on the manuscript. This work was supported by grants from the National Institutes of Health.

Author Contributions C.T.M. and E.E.F. conceived and designed the experiments, C.T.M. and S.I.P. performed the experiments, and C.T.M. and E.E.F. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.T.M. (ctmoritz@u.washington.edu).

METHODS

Cortical recording. Sterile surgeries were performed with isoflurane anaesthesia (1–1.5% in 50:50 O₂:N₂O). All surgeries were followed by a program of analgesics (buprenorphine 0.15 mg kg⁻¹ and ketoprofen 5 mg kg⁻¹) and antibiotics (cephalexin 25 mg kg⁻¹). Each animal was implanted with a cranial recording chamber over the left hand and wrist area of the motor cortex using stereotaxic coordinates (anterior: 13 mm, lateral: 18 mm) to permit cortical recordings^{29,30}. To obtain longer duration cell recordings, monkey L was re-implanted with a chronic electrode array over the left motor cortex. The array of 12 independently movable microwires is fully described elsewhere³¹. In brief, 50 µm tungsten wires were threaded through individual polyamide guide-tubes in a 2 × 6 array that was anchored to the skull. This array provided stable recordings from the same isolated cell for the duration of an experimental session, and across multiple days for ten cells^{24,31,32}.

Nerve block implant. Reversible paralysis of the right wrist was achieved with one of two nerve block methods. First, catheters were implanted in the brachial plexus, near cords giving rise to the radial, ulnar and median nerves. Epidural catheters (19 gauge, Arrow International) were inserted into the epineurium surrounding each nerve and anchored in place with cyanoacrylate. Second, nerve cuffs with catheters³³ were implanted around the median and ulnar nerves in the upper arm. Catheters terminating in the lumen of each Silastic cuff (4 mm inner diameter, 30 mm long) permitted the nerves to be bathed in anaesthetic. Nerves were identified by electrical stimulation, and catheters were tunnelled subcutaneously to exit the skin on the upper back and sealed with an injection port. Thirty-one cells controlled FES during nerve blocks induced by the catheter method, and the remaining 13 cells during blocks induced by cuffs.

Experimental procedures. The monkey sat with his right elbow and hand immobilized by padded splints while a transducer measured the flexion-extension (F-E) and radial-ulnar (R-U) torques produced about the wrist (see Fig. 1a). To receive an applause reward, the monkey maintained wrist torque within a centre target (zero torque) followed by one of eight peripheral targets specifying different combinations of F-E and R-U torque (average magnitude 0.13 ± 0.01 nM). Isolated cell activity was discriminated online using template-matching software (Alpha Omega MSD). Subsequently, cell activity controlled cursor movement in one dimension. Interspike intervals were averaged over a 0.5-s sliding window to create a continuous signal for cursor position (and later FES control). If the cell was directionally tuned, targets were aligned with its preferred direction. For untuned cells or cells without tuning information (that is, cells isolated after nerve block began), either the left or right screen position was arbitrarily chosen to represent high discharge rates for visual feedback. Monkeys practiced cell control for 10 min, maintaining discharge rates within each target for 0.5–1.0 s to receive a reward.

Nerve block. We blocked nerves leading to wrist muscles with local anaesthetic to create temporary motor paralysis. Block onset typically occurred after 5–60 min, depending on anaesthetic and block method. During this time the monkeys continued to perform the cell-controlled target-tracking task. Further doses were given regularly to maintain paralysis during FES control.

Brain-controlled FES. After onset of paralysis and an average of 36 ± 22 min of cell-controlled target tracking, the cell activity was converted to stimuli delivered to one or more paralysed muscles. Wrist torque controlled the position of the cursor, and targets were randomly displayed on the monitor in one dimension. Monkeys were required to maintain torque within each target for 0.5–1.0 s (mean 0.56 s) to receive a reward. Targets remained on the screen until satisfied, followed by presentation of the next target either immediately or after a variable

reward period (1.5–1.7 s). Forty-two cells controlled stimulation current in proportion to cell rate, permitting the monkey to grade contraction force. Two of these cells also controlled stimulation via the autonomous 'Neurochip'^{23,24} to deliver a 1-s train of stimuli (2.5 mA, 50 s⁻¹) when smoothed cell rate exceeded a threshold (Supplementary Fig. 4). Similarly, the first two cells in monkey I also triggered a 1-s train of 50 s⁻¹ stimuli at 5 mA.

To confirm continued nerve block during the practice session, the stimulator was turned off after every 10 min of FES to assure that the monkey could not acquire the peripheral target through volitional muscle contractions. Figure 2b and Supplementary Fig. 4 illustrate the torques produced with the stimulator active compared to periods when the stimulator was turned off for 30 s. With the stimulator off, the monkey repeatedly attempted to satisfy the target but produced $\leq 10\%$ of the torque used to acquire the active target. For all such test periods, with each cell the monkeys produced an average maximum of only $18.0 \pm 21.3\%$ of the torque used to satisfy the targets.

Data sampling. Signals were digitized and stored to disk for offline analysis. Raw recording from motor cortex was band-passed from 1–10 kHz and sampled at 25 kHz, along with spike times from the online discrimination. Wrist torques (flexion-extension and radial-ulnar) were sampled at 5 kHz, and smoothed and down-sampled to 500 Hz during offline analysis. We also recorded behavioural parameters (for example, target on screen), and muscle stimulation amplitude and timing (5 kHz).

Data analysis. Task difficulty was increased incrementally by raising levels of torque targets and increasing hold times. This complicated the quantification of skill learning. Improvements were evident as higher performance levels before increments in task requirements (see, for example, Supplementary Fig. 3), and these times were compared with performance at the beginning of a practice session. Specifically, the two-minute period with the peak performance was compared to the initial two minutes of practice (for example, targets per minute). Control precision was measured by target errors and the number of failed attempts to reach a target. A target acquisition error was counted when the monkey activated the stimulator while the centre target was on the screen, resulting in sufficient torque to satisfy any peripheral target had it been presented. Target errors are reported as the percentage of centre targets presented. A failed attempt was counted whenever the monkey briefly acquired a peripheral torque target but did not satisfy the required hold time. A Student's *t*-test was used to compare average torques during graded FES control. Otherwise, the non-parametric rank sum test was used for all comparisons as at least one data set for each remaining comparison failed the Lilliefors test for normality. All reported values are means \pm s.d.

29. Evarts, E. V. Relation of pyramidal tract activity to force exerted during voluntary movement. *J. Neurophysiol.* **31**, 14–27 (1968).
30. Woolsey, C. N. et al. Patterns of localization in precentral and "supplementary" motor areas and their relation to the concept of a premotor area. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* **30**, 238–264 (1952).
31. Jackson, A. & Fetz, E. E. A compact moveable microwire array for long-term chronic unit recording in cerebral cortex of primates. *J. Neurophysiol.* **98**, 3109–3118 (2007).
32. Jackson, A., Mavoori, J. & Fetz, E. E. Correlations between the same motor cortex cells and arm muscles during a trained task, free behavior, and natural sleep in the macaque monkey. *J. Neurophysiol.* **97**, 360–374 (2007).
33. Loeb, G. E. & Hoffer, J. A. Activity of spindle afferents from cat anterior thigh muscles. II. Effects of fusimotor blockade. *J. Neurophysiol.* **54**, 565–577 (1985).

Sox18 induces development of the lymphatic vasculature in mice

Mathias François¹, Andrea Caprini², Brett Hosking¹, Fabrizio Orsenigo², Dagmar Wilhelm¹, Catherine Browne¹, Karri Paavonen³, Tara Karnezis³, Ramin Shayan^{3,4}, Meredith Downes¹, Tara Davidson¹, Desmond Tutt¹, Kathryn S. E. Cheah⁵, Steven A. Stacker³, George E. O. Muscat¹, Marc G. Achen³, Elisabetta Dejana^{2,6} & Peter Koopman¹

The lymphatic system plays a key role in tissue fluid regulation and tumour metastasis, and lymphatic defects underlie many pathological states including lymphoedema, lymphangiectasia, lymphangioma and lymphatic dysplasia^{1–3}. However, the origins of the lymphatic system in the embryo, and the mechanisms that direct growth of the network of lymphatic vessels, remain unclear. Lymphatic vessels are thought to arise from endothelial precursor cells budding from the cardinal vein under the influence of the lymphatic hallmark gene *Prox1* (prospero homeobox 1; ref. 4). Defects in the transcription factor gene *SOX18* (SRY (sex determining region Y) box 18) cause lymphatic dysfunction in the human syndrome hypotrichosis-lymphoedema-telangiectasia⁵, suggesting that Sox18 may also play a role in lymphatic development or function. Here we use molecular, cellular and genetic assays in mice to show that Sox18 acts as a molecular switch to induce differentiation of lymphatic endothelial cells. Sox18 is expressed in a subset of cardinal vein cells that later co-express *Prox1* and migrate to form lymphatic vessels. Sox18 directly activates *Prox1* transcription by binding to its proximal promoter. Overexpression of Sox18 in blood vascular endothelial cells induces them to express *Prox1* and other lymphatic endothelial markers, while *Sox18*-null embryos show a complete blockade of lymphatic endothelial cell differentiation from the cardinal vein. Our findings demonstrate a critical role for Sox18 in developmental lymphangiogenesis, and suggest new avenues to investigate for therapeutic management of human lymphangiopathies.

Sox18 belongs to the SRY-related HMG domain family of developmental transcription factors^{6,7}. During embryogenesis, *Sox18* is expressed in developing vasculature and hair follicles⁸, consistent with the reported cardiovascular and coat texture phenotype of naturally occurring mouse mutants of the *ragged* (*Ra*) allelic series^{9,10}. The phenotype of *ragged* mutants is evidently due to dominant-negative action of the mutant Sox18 protein (Sox18^{Ra}) produced in these mice¹¹, resulting from the normal ability of Sox18^{Ra} to occupy binding sites in the promoters of Sox18 target genes, but a failure to engage partners and/or cofactors required to initiate transcription of these targets¹².

To explore the role of Sox18 in lymphatic vascular development, we first re-examined the phenotype of *ragged-opossum* (*Ra*^{Op}), the most severely affected class of *Ra* mutant^{11,13}. Most homozygous *Ra*^{Op}/*Ra*^{Op} embryos showed gross subcutaneous oedema at 13.5 days post coitum (d.p.c.; Fig. 1a). We also bred mice with a targeted inactivation ('knockout') of *Sox18* from their original mixed 129-CD1

background¹⁴ onto a pure C57BL/6 (B6) background by 11 generations of backcrossing to inbred B6 mice; the resulting strain, designated *Sox18*^{-/-} (B6), was used throughout this study. This backcrossing revealed gross subcutaneous oedema (Fig. 1a) and fetal lethality in all homozygotes, and mild subcutaneous oedema around the head in some heterozygotes (2 of 4; not shown), features not evident on the mixed background.

Ra^{Op}/*Ra*^{Op} and *Sox18*^{-/-} (B6) embryos die shortly after 14.5 d.p.c., precluding study of lymphatic physiology beyond that stage. We therefore examined the structure of the mature lymphatic vascular network in adult mice with heterozygous *Sox18* mutations (*Ra*^{Op}/+ and *Sox18*^{+/-} B6), by immunofluorescence analysis of whole ear tissue, using antibodies specific for the marker LYVE-1. The lymphatic vessels in *Sox18*-mutant mice were finer, denser and more branched, with more loops and fewer blind-ended sacs, compared to wild-type mice (Fig. 1b, c). In contrast, blood vascular patterning was unaffected in *Ra*^{Op}/+ and *Sox18*^{+/-} (B6) mice, as revealed by PECAM-1 staining (Fig. 1b). Thus, defects in *Sox18* function are sufficient to alter the patterning of the lymphatic vasculature, and imply that lymphatic dysgenesis contributes to the oedema seen in the *Sox18* mutant mice and in hypotrichosis-lymphoedema-telangiectasia (HLT).

We next examined the timing and location of Sox18 expression during mouse embryo development, with reference to the known molecular and cellular hallmarks of developmental lymphangiogenesis. At 9 d.p.c., expression of Sox18, but not *Prox1*, was observed in endothelial cells lining the dorsolateral sector of the cardinal vein (Fig. 2a, arrows). At 10.5 d.p.c., Sox18 expression continued in a population of cells migrating from this area; these cells also expressed *Prox1* and the endothelial marker PECAM-1, indicating that they are precursors of the lymphatic vasculature¹⁵ (Fig. 2b, c, yellow nuclei). Sox18 continued to be expressed in most cells of the primary lymph sacs at 13.5 d.p.c. (Fig. 2d, e, arrows). These cells also expressed PECAM-1 and the lymphatic marker podoplanin, indicating their lymphatic nature. At this stage, Sox18 expression was also maintained in some blood endothelial cells, expressing PECAM-1 but not podoplanin, outside the lymph sacs (Fig. 2f, asterisks). By 14.5 d.p.c., Sox18 expression had subsided in the lymph sacs (Fig. 2g), but could still be detected in mesenteric and dermal blood vessels (Fig. 2h, i, red). Sox18 was not detectable in mesenteric or dermal lymphatics during their early development (Fig. 2h, i, arrows) or in newborn mice (Fig. 2j, k), or in mature lymphatic vessels in the adult (Fig. 2l), signifying that this protein is not required for maintenance of lymphatic endothelial cells once they have differentiated.

¹Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia. ²IFOM, FIRC Institute of Molecular Oncology, Via Adamello 16, 20139 Milan, Italy. ³Ludwig Institute for Cancer Research, PO Box 2008, Royal Melbourne Hospital, Victoria 3050, Australia. ⁴Department of Surgery, The University of Melbourne, Parkville, Victoria 3052, Australia. ⁵Department of Biochemistry and Centre for Reproduction, Development & Growth, University of Hong Kong, Pokfulam, Hong Kong, China. ⁶Department of Biomolecular Sciences and Biotechnologies, University of Milan, 20129 Milan, Italy.

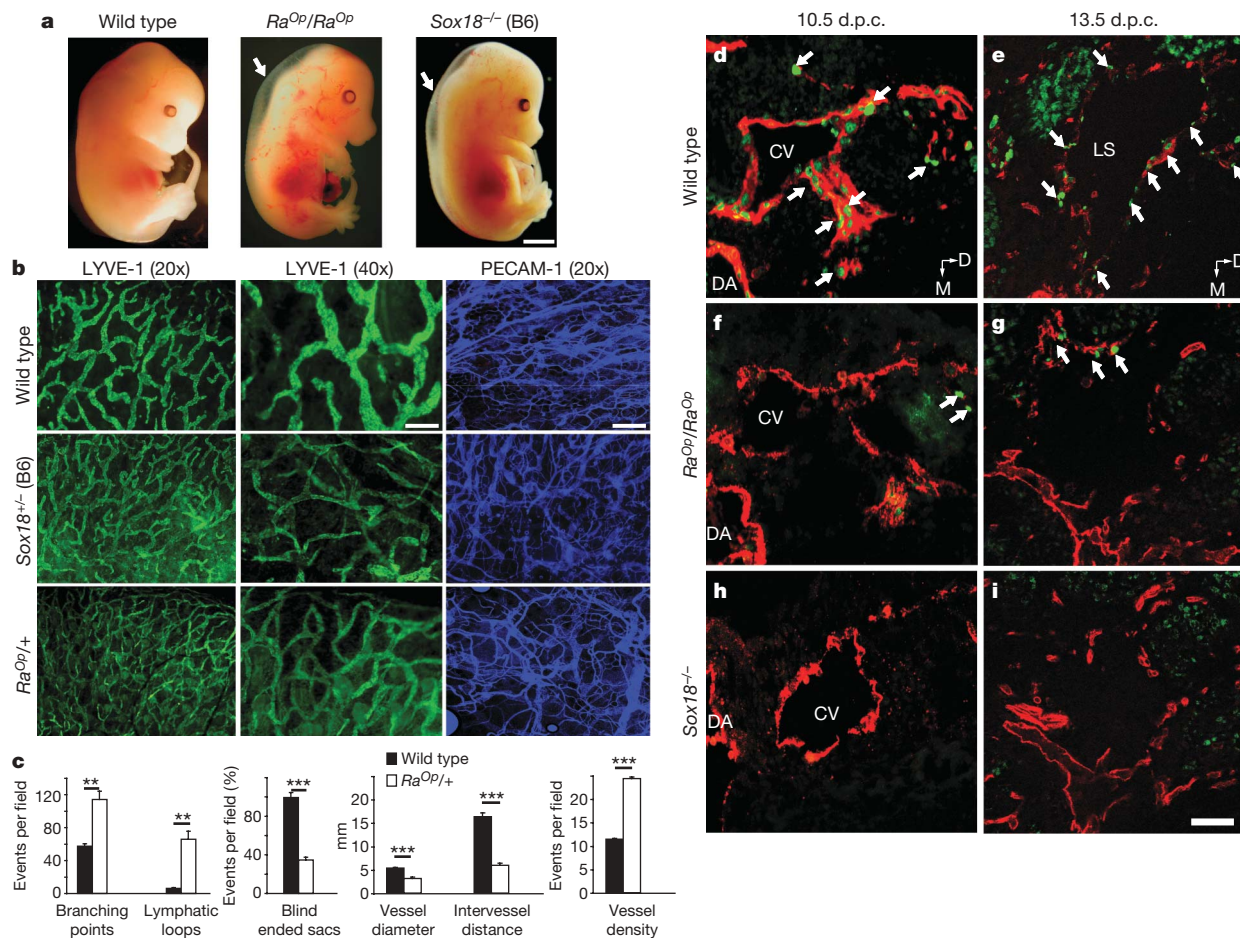


Figure 1 | Lymphatic vascular defects resulting from *Sox18* dysfunction. **a**, *Sox18*^{−/−} (B6) and *Ra*^{Op}/*Ra*^{Op} mutant embryos show oedema at 13.5 d.p.c. (arrows). **b**, **c**, Lymphatic defects in adult *Sox18* mutant mice. **b**, Immunofluorescence for the lymphatic marker LYVE-1 (green) on whole ear segments revealed clear differences between the large-diameter, blunt-ended lymphatic vessels in wild-type (upper row) and the smaller, highly branched vessels in *Sox18*^{+/−} (B6) (middle row) and *Ra*^{Op}/*+* mice (lower row). Immunostaining for the general endothelial cell marker PECAM-1 (blue) showed a similar patterning of the blood vasculature in both types of mice. **c**, Quantitative analysis of lymphatic parameters in *Ra*^{Op}/*+* versus wild-type ears based on LYVE-1 immunostaining. *Ra*^{Op}/*+* heterozygotes showed a statistically significant difference in all parameters observed compared to wild type. The averaged field counts for each parameter were

collated for all sections and their respective statistical significance determined by Student's *t*-test ($n = 3$). Data represent mean + s.e.m.; *** $P < 0.001$, ** $P < 0.01$. **d–i**, Loss of lymphatic endothelial differentiation in *Sox18*-mutant embryos. Prox1 (green)/PECAM-1 (red) double immunofluorescence on transverse sections. **d**, At 10.5 d.p.c. and **e**, 13.5 d.p.c., Prox1/PECAM-1-positive lymphatic endothelial cells can be seen in the cardinal vein and jugular lymph sac areas in wild-type embryos (arrows). **f**, **g**, Markedly fewer Prox1/PECAM-1-positive cells were detected in *Ra*^{Op}/*Ra*^{Op} embryos (arrows). **h**, **i**, No Prox1-positive lymphatic endothelial cells could be detected at all in *Sox18*-null embryos. CV, cardinal vein; DA, dorsal aorta; LS, lymph sac; D, dorsal, M, medial. Scale bars, 1.6 mm (a); 40 μ m (b, 20 \times); 20 μ m (b, 40 \times , d–i).

To investigate whether *Sox18* is able to modulate the expression of *Prox1* and other lymphatic markers in an *in vitro* model system, we overexpressed wild-type *Sox18* or dominant-negative *Sox18*^{RaOp} in mouse embryonic stem (ES) cells by lentiviral infection and examined the resulting expression of lymphatic marker genes. In the system studied, the ES cells form embryoid bodies and spontaneously differentiate into endothelial cells that express *Prox1* and *podoplanin*^{6,17}. Expression of both markers was significantly increased through overexpression of *Sox18*, while overexpression of *Sox18*^{RaOp} strongly inhibited induction of lymphatic markers (Fig. 3a). Moreover, the inhibition of *Prox1* and *podoplanin* expression by *Sox18*^{RaOp} was rescued by infection with high levels of *Sox18*-expressing lentivirus (Fig. 3a), indicating that the effects of *Sox18*^{RaOp} were mediated by direct interference with *Sox18* function in this system. Overexpression of *Sox18* and *Sox18*^{RaOp} did not affect general blood vascular endothelial cell differentiation markers such as *Vegfr2*, *Tie2* and *VE-cadherin* (Fig. 3b). Therefore, *Sox18* stimulates expression of lymphatic markers during the differentiation of ES cells along an endothelial-specific pathway.

To further investigate the possible role of *Sox18* at the nexus of vascular and lymphatic endothelial cell specification, we used an

additional cell-based model system. H5V mouse blood endothelial cells¹⁸ were infected with the *Sox18*, *Sox18*^{RaOp} or GFP-expressing lentiviruses described above, and lymphatic marker expression was analysed by quantitative RT-PCR. *Sox18* overexpression in these cells resulted in significant upregulation of the lymphatic markers *Prox1*, *ephrinB2* and *Vegfr3* (Fig. 3c). Enhanced *Prox1* protein expression in nuclei of H5V cells after *Sox18* overexpression was confirmed by immunofluorescence (Supplementary Fig. 1). In contrast, cells overexpressing the mutant *Sox18*^{RaOp} protein did not show any significant change in the expression level of these lymphatic-specific markers relative to GFP-infected cells (Fig. 3c), probably because H5V cells express only basal levels of lymphatic markers that cannot be reduced further by suppressing *Sox* function. The induction of lymphatic markers in this assay demonstrates that *Sox18* overexpression causes blood vascular endothelial cells to activate the lymphatic endothelial cell pathway. In addition, levels of the blood vascular markers N-cadherin and CD44, but not *Nrp1*, *endoglin* or *cd34*, were reduced after *Sox18* overexpression (Supplementary Fig. 2), supporting the concept that *Sox18*, acting via *Prox1*, downregulates at least some blood endothelial markers while promoting the lymphatic endothelial differentiation program.

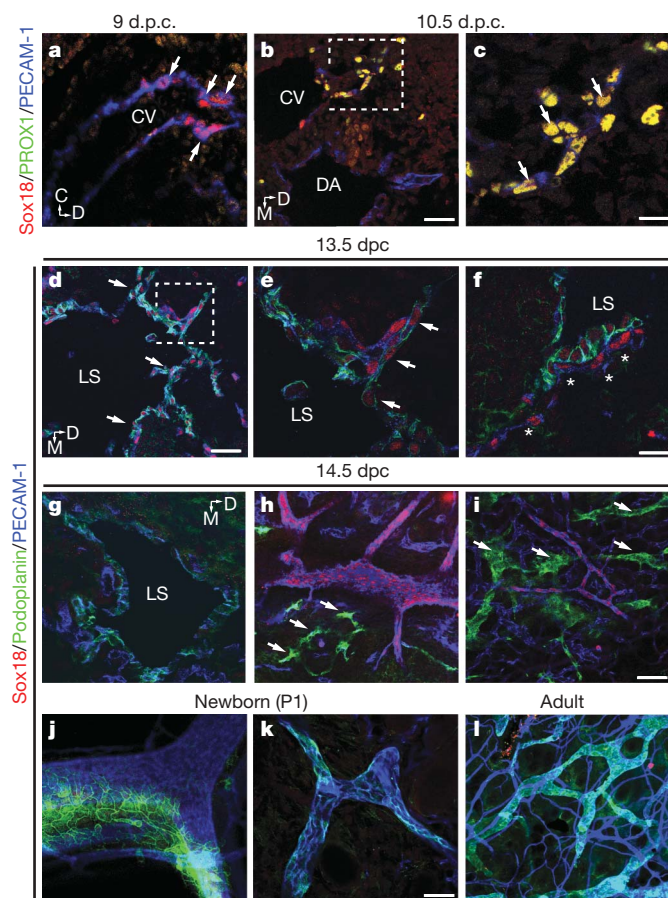


Figure 2 | Sox18 expression during lymphatic vascular development. **a–c**, Immunofluorescence for Sox18 (red), endothelial marker PECAM-1 (blue) and Prox1 (green). Lymphatic endothelial progenitors (yellow) co-express Sox18 and Prox1 at 10.5 d.p.c. (**c**, boxed area in **b**, arrows). **d–f**, Immunofluorescence for Sox18 (red), endothelial marker PECAM-1 (blue) and lymphatic marker podoplanin (green). **e**, Enlargement of boxed area in **d**, arrows indicate triple-positive lymphatic endothelial cells in lymph sacs at 13.5 d.p.c. **f**, Some blood endothelial cells expressing Sox18 and PECAM-1 but not podoplanin are visible outside the lymph sacs (asterisks). **g**, Lack of Sox18 expression in lymph sacs, mesenteric lymphatics (**h**, arrows) and dermal lymphatics (**i**, arrows) by 14.5 d.p.c. **j**, Newborn mesentery, **k**, newborn skin and **l**, mature lymphatic vessels at adulthood, lack Sox18 expression. C, caudal V, ventral; M, medial; D, dorsal; CV, cardinal vein; LS, lymph sac; DA, dorsal aorta. Scale bars: 50 μ m (**a**, **b**, **d**, **g**); 20 μ m (**j**, **k**, **l**); 10 μ m (**c**, **e**, **f**, **h**, **i**).

To determine whether Sox18 regulates *Prox1* expression directly, we cloned 3,952 base pairs (bp) of the *Prox1* promoter and identified two putative Sox binding sites at $-1,135$ to $-1,130$ bp (*SoxA*) and -813 to -808 bp (*SoxB*) relative to the translation start site. We tested the binding of Sox18 to these putative response elements by electrophoretic mobility shift assays (EMSA). Bacterially expressed, purified GST–Sox18 fusion protein bound strongly to both *SoxA* and *SoxB* sites (Fig. 4a), while the GST-only control protein bound neither site. Specificity of binding was confirmed by competing the interaction with increasing amounts of unlabelled oligonucleotide. Sox18^{RaOp} was able to bind to the same sites in the *Prox1* promoter (Fig. 4a) with a similar binding affinity to that of wild-type Sox18, supporting previous evidence that *ragged* forms of Sox18 act in a dominant-negative fashion by binding to Sox target sequences but failing to activate transcription of target genes^{11,19}.

As an independent confirmation of these data, chromatin immunoprecipitation (ChIP) experiments were performed using chromatin from the mouse mesenteric lymph node endothelial cell line mEnd (ref. 20), which expresses endogenous Sox18 and Prox1 (Supplementary Fig. 3). DNA fragments precipitated by anti-Sox18

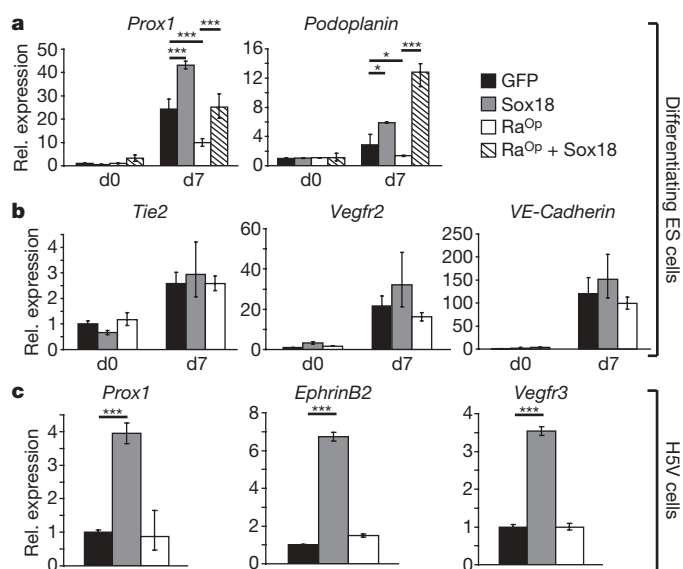


Figure 3 | Sox18 induces lymphatic markers in differentiating ES cells and blood vascular endothelial cells. **a**, ES cells were infected with lentivirus expressing GFP (an inert negative control; black bars), Sox18 (grey bars), Sox18^{RaOp} (white bars), or Sox18^{RaOp} together with an excess of Sox18 (hatched bars) and induced to differentiate *in vitro*. Control cells expressed the lymphatic markers *Prox1* and *podoplanin*, as assessed by quantitative RT-PCR, within 7 days of culture. Both markers were further upregulated by Sox18 and downregulated by Sox18^{RaOp}; the effect of Sox18^{RaOp} could be rescued by overexpression of Sox18. **b**, Sox18 or Sox18^{RaOp} overexpression did not affect blood vascular endothelial cell differentiation markers *Tie2*, *Vegfr2* or *VE-Cadherin*. **c**, H5V blood vascular endothelial cells were infected with lentivirus expressing GFP (black bars), Sox18 (grey bars) or Sox18^{RaOp} (white bars) and assayed for marker gene expression by quantitative PCR. Lymphatic markers *Prox1*, *EphrinB2* and *Vegfr3* were induced by Sox18 but not by Sox18^{RaOp}. For each set of studies the figure represents a typical experiment of at least three performed. Data are means \pm s.d. of six replicates and are expressed relative to GFP. *** $P < 0.01$, * $P < 0.05$ versus GFP expressing cells by Student's *t*-test.

antibody were purified and tested for the presence of the *SoxA/B* binding sites in the *Prox1* promoter by PCR-amplifying a 521 bp fragment. After 22 PCR cycles, a clear enrichment of this fragment was detected (Fig. 4b), confirming that Sox18 protein binds to the *Prox1* promoter in mammalian cells.

We next determined whether Sox18 could *trans*-activate the *Prox1* promoter driving a luciferase reporter gene after transfection into mEnd cells. Overexpression of Sox18 led to a significant increase in luciferase activity (Fig. 4c). This activation was observed using the original 4 kb *Prox1* promoter fragment that contains both *SoxA* and *SoxB* binding sites. Overexpression of the mutant Sox18^{RaOp} did not enhance the luciferase activity driven by the *Prox1* promoter (Fig. 4c). Site directed mutagenesis of the *Prox1* promoter, which altered key sequence elements of either the *SoxA* site or the *SoxB* site, abolished the enhanced activity produced by Sox18 overexpression (Fig. 4c). Hence, activation of the *Prox1* promoter in this assay system requires cooperation of the identified *SoxA* and *SoxB* sites.

To confirm the importance of the 4 kb promoter fragment in the *in vivo* regulation of *Prox1* expression, we generated transgenic mice carrying a GFP reporter construct driven by this fragment. In transgenic embryos at 10.5 d.p.c., GFP-expressing cells were found in the region of the cardinal vein corresponding to the location of the Prox1-positive lymphatic progenitors at that stage (Fig. 4d, Supplementary Fig. 4a, arrows). By 16.5 d.p.c., a complex network of GFP-positive vessels was observed, identical to the pattern of endogenous Prox1 expression in the lymphatic vascular network (Supplementary Fig. 4b). Co-localization of GFP with the lymphatic marker podoplanin in transgenic embryos confirmed transgene expression in lymphatic endothelial cells in these mice (Fig. 4e).

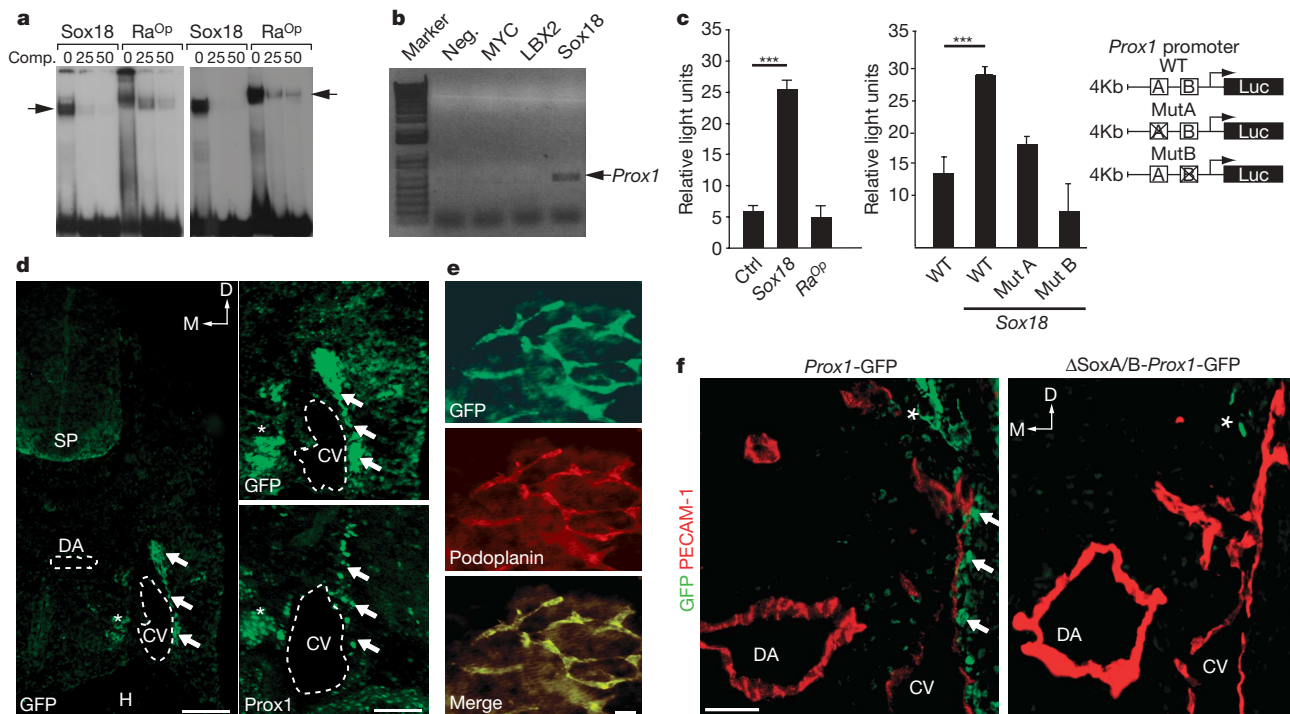


Figure 4 | Sox18 activates the *Prox1* promoter *in vitro* and *in vivo*. **a**, EMSA showing binding of Sox18 and Sox18^{RaOp} to SoxA (left panel) and SoxB (right panel) binding sites in the *Prox1* promoter (arrows). Specificity was confirmed using a 25- or 50-fold molar excess of non-radioactive competitor oligonucleotide (Comp.). **b**, Chromatin immunoprecipitation (ChIP) of endogenous Sox18 in mEnd cells. Negative controls omitting PCR template (Neg.), or using anti-MYC or anti-LBX2 antibodies are shown. *Prox1* fragment is arrowed. **c**, Reporter gene assays using constructs containing either wild-type *Prox1* promoter (left graph) or versions with the corresponding Sox responsive element mutated (right graph). Data are means \pm s.e.m. of three independent experiments each performed in triplicate; *** P < 0.001. **d**, *In vivo* analysis of the 4 kb *Prox1* promoter-GFP construct in transgenic embryos. GFP signal (left and upper right panels, arrows) was detected in cells adjacent to the cardinal vein in 10.5 d.p.c.

These observations indicate that the Sox18-responsive 4 kb *Prox1* promoter fragment identified in this study is sufficient to drive faithful expression of GFP during developmental lymphangiogenesis *in vivo*. This promoter fragment also drove reporter gene expression in other tissue types such as the sympathetic ganglia, where endogenous *Prox1* is also normally expressed²¹ (Fig. 4d, asterisks). This fragment evidently contains elements required for the expression of *Prox1* in those sites and also for maintenance of lymphatic endothelial expression of *Prox1* after Sox18 expression has subsided. Ablation of the SoxA and SoxB sites in the transgene abolished *Prox1* promoter activity specifically in the cardinal vein area where lymphatic endothelial precursor cells are generated (Fig. 4f, arrows), but not in the sympathetic ganglia (Fig. 4f, asterisk), indicating that Sox18 binding to the *Prox1* promoter is specifically required in the context of lymphangiogenesis *in vivo*.

Finally, to test whether Sox18 is also necessary for stimulating *Prox1* expression in lymphatic precursor cells *in vivo*, we examined the expression pattern of *Prox1* in Sox18-mutant embryos, in comparison to their wild-type littermates. At 10.5 d.p.c., when extensive migration of *Prox1*-expressing lymphatic endothelial precursor cells normally occurs in the mouse embryo (Fig. 1d), the number of *Prox1*-positive cells within and around the cardinal vein was greatly reduced in *RaOp/RaOp* embryos (Fig. 1f). Moreover, these cells were completely absent in Sox18-null (B6) embryos at this stage (Fig. 1h, Supplementary Fig. 5). This total lack of lymphatic endothelial cells was confirmed by complete serial section analysis of 10.5 d.p.c. Sox18-null (B6) embryos. Expression of *Prox1* in non-lymphatic cell

embryos. This expression pattern coincided with endogenous *Prox1* expression at 10.5 d.p.c. (lower right panel, arrows). GFP and *Prox1* expression were also observed in the sympathetic ganglia (asterisks). **e**, GFP expression in the lymphatic network of the skin of 16.5 d.p.c. transgenic embryos (top panel), co-incident with podoplanin (centre panel; merged pictures, bottom panel). **f**, Genetic ablation of SoxA and SoxB sites by site-directed mutagenesis (Δ SoxA/B-*Prox1*-GFP) abolishes *Prox1* promoter activity *in vivo*. At 10.5 d.p.c., transgenic embryos generated with the wild-type *Prox1*-promoter display GFP expression in PECAM-1 positive cells migrating from the dorsal part of the cardinal vein (left, arrows), whereas transgenic embryos (right) lack these cells. Asterisk indicates residual GFP expression in PECAM-1-negative neural cells. DA, dorsal aorta; CV, cardinal vein; D, dorsal; M, medial. Scale bars: 100 μ m (**d**, left panel); 50 μ m (**d**, right panels, **f**); 10 μ m (**e**).

types such as the myocardium was not affected by the lack of Sox18 (Supplementary Fig. 5). The blockade of lymphatic development cannot be ascribed to a primary deficiency in blood vascular endothelial cells; no defects in blood vascular structure or marker expression were observed in the area of the cardinal vein, or anywhere else in mutant embryos (Supplementary Fig. 6 and data not shown).

By 13.5 d.p.c., linear groups of *Prox1*-positive cells were clearly visible lining the primary lymph sacs in wild-type embryos (Fig. 1e), as described previously⁴. Similar cells were present in only low numbers in *RaOp/RaOp* embryos (Fig. 1g), and completely absent in Sox18-null embryos (Fig. 1i). Both *RaOp/RaOp* and Sox18-null embryos exhibited a hypertrophic jugular vein and absence of lymph sacs (Supplementary Fig. 7). These observations establish that lack of Sox18 function results in a complete blockade of *Prox1*-mediated lymphangiogenesis from the pool of vascular endothelial cell precursors during mouse embryogenesis, confirming the pivotal role of Sox18 in this system.

In summary, our data suggest a model in which Sox18 acts at the nexus of differentiation of lymphatic endothelial progenitor cells from blood vascular precursors in the embryo (Supplementary Fig. 8). The role of Sox18 in promoting lymphatic endothelial cell specification is consistent with the known roles of other Sox transcription factors that act as developmental switches^{6,22}. Moreover, Sox18, like other Sox factors, clearly acts in a number of developmental systems—hair follicle development, vasculogenesis and lymphangiogenesis—and is likely to depend on context-specific partner proteins and/or cofactors^{23,24}. Further studies will focus on defining the molecular milieu

in which Sox18 operates to activate *Prox1* transcription, and may reveal new strategies for therapeutic stimulation or suppression of lymphangiogenesis.

METHODS SUMMARY

Antibodies were obtained commercially, except anti-Sox18 (Supplementary Fig. 9), which was raised in rabbits against amino acids (a.a.) 228–254 of mouse Sox18 and biotinylated as described²⁵. For transgenesis, CBA/C57BL6 F1 zygotes were injected with *Prox1*-GFP and reimplanted into outbred foster mothers. EMSA was carried out as described¹⁹. For ChIP, *Prox1* primers 5'-GGCAAGGGTGTTCCTTTT-3' and 5'-GGCACAGACATGCACTGTCC-3' were used with 50 °C annealing and 22 PCR cycles.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 December 2007; accepted 5 September 2008.

Published online 19 October 2008.

- Adams, R. H. & Alitalo, K. Molecular regulation of angiogenesis and lymphangiogenesis. *Nature Rev. Mol. Cell Biol.* **8**, 464–478 (2007).
- Carmeliet, P. Angiogenesis in health and disease. *Nature Med.* **9**, 653–660 (2003).
- Stacker, S. A. *et al.* Lymphangiogenesis and cancer metastasis. *Nature Rev. Cancer* **2**, 573–583 (2002).
- Wigle, J. T. *et al.* An essential role for Prox1 in the induction of the lymphatic endothelial cell phenotype. *EMBO J.* **21**, 1505–1513 (2002).
- Irrthum, A. *et al.* Mutations in the transcription factor gene SOX18 underlie recessive and dominant forms of hypotrichosis-lymphedema-telangiectasia. *Am. J. Hum. Genet.* **72**, 1470–1478 (2003).
- Bowles, J., Schepers, G. & Koopman, P. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev. Biol.* **227**, 239–255 (2000).
- Hosking, B. M. *et al.* Cloning and functional analysis of the Sry-related HMG box gene, Sox18. *Gene* **262**, 239–247 (2001).
- Pennisi, D. *et al.* Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice. *Nature Genet.* **24**, 434–437 (2000).
- Carter, T. C. & Philipps, R. J. S. Ragged, a semi-dominant coat texture mutant in the house mouse. *J. Hered.* **45**, 151–154 (1954).
- Slee, J. The morphology and development of ragged — a mutant affecting the skin and hair of the house mouse. II. Genetics, embryology and gross juvenile morphology. *J. Genet.* **55**, 570–584 (1957).
- James, K. *et al.* Sox18 mutations in the ragged mouse alleles ragged-like and opossum. *Genesis* **36**, 1–6 (2003).
- Hosking, B. M. *et al.* SOX18 directly interacts with MEF2C in endothelial cells. *Biochem. Biophys. Res. Commun.* **287**, 493–500 (2001).
- Green, E. L. & Mann, S. J. Opossum, a semi-dominant lethal mutation affecting hair and other characteristics of mice. *J. Hered.* **52**, 223–227 (1961).
- Pennisi, D. *et al.* Mice null for Sox18 are viable and display a mild coat defect. *Mol. Cell. Biol.* **20**, 9331–9336 (2000).
- Wigle, J. T. & Oliver, G. Prox1 function is required for the development of the murine lymphatic system. *Cell* **98**, 769–778 (1999).
- Balconi, G., Spagnuolo, R. & Dejana, E. Development of endothelial cell lines from embryonic stem cells: A tool for studying genetically manipulated endothelial cells *in vitro*. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1443–1451 (2000).
- Liersch, R., Nay, F., Lu, L. & Detmar, M. Induction of lymphatic endothelial cell differentiation in embryoid bodies. *Blood* **107**, 1214–1216 (2006).
- Garlanda, C. *et al.* Progressive growth in immunodeficient mice and host cell recruitment by mouse endothelial cells transformed by polyoma middle-sized T antigen: Implications for the pathogenesis of opportunistic vascular tumors. *Proc. Natl Acad. Sci. USA* **91**, 7291–7295 (1994).
- Hosking, B. M. *et al.* The VCAM-1 gene that encodes the vascular cell adhesion molecule is a target of the Sry-related high mobility group box gene, Sox18. *J. Biol. Chem.* **279**, 5314–5322 (2004).
- Sorokin, L. *et al.* Expression of novel 400-kDa laminin chains by mouse and bovine endothelial cells. *Eur. J. Biochem.* **223**, 603–610 (1994).
- Rodriguez-Niedenführ, M. *et al.* Prox1 is a marker of ectodermal placodes, endodermal compartments, lymphatic endothelium and lymphangioblasts. *Anat. Embryol. (Berl.)* **204**, 399–406 (2001).
- Wegner, M. From head to toes: The multiple facets of Sox proteins. *Nucleic Acids Res.* **27**, 1409–1420 (1999).
- Kamachi, Y., Uchikawa, M. & Kondoh, H. Pairing SOX off with partners in the regulation of embryonic development. *Trends Genet.* **16**, 182–187 (2000).
- Wilson, M. & Koopman, P. Matching SOX: Partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr. Opin. Genet. Dev.* **12**, 441–446 (2002).
- Wilhelm, D. *et al.* Sertoli cell differentiation is induced both cell-autonomously and through prostaglandin signaling during mammalian sex determination. *Dev. Biol.* **287**, 111–124 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Nagy, W. Abramow-Newerly, L. Naldini, P. Berta, A. Yap and J. Gamble for gifts of reagents. We also thank L. Bernard, L. Tizzoni and V. Dall'Olio for quantitative real-time PCR analysis, M. Chan and A. Pelling for antibody characterization, and S. Pizzi for technical assistance. Confocal microscopy was performed at the Australian Cancer Research Foundation Dynamic Imaging Centre for Cancer Biology. This work was supported by the National Health and Medical Research Council (Australia), INSERM (France), the Associazione Italiana per la Ricerca sul Cancro (Italy), the European Community (LSHG-CT-2004-503573, Eustroke and Optistem networks), the Pfizer Foundation (Australia), the Research Grants Council (Hong Kong), the Raelene Boyle Sporting Chance Foundation (Australia), the Royal Australasian College of Surgeons, the Heart Foundation of Australia, and the Australian Research Council.

Author Contributions M.F., A.C., B.H., F.O., D.W., C.B., K.P., T.K., R.S., M.D., T.D. and D.T. conducted the experiments, and K.S.E.C., S.A.S., G.E.O.M., M.G.A., E.D. and P.K. supervised the work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.K. (p.koopman@imb.uq.edu.au) or E.D. (elisabetta.dejana@ifom-ieu-campus.it).

METHODS

Mouse strains. Heterozygous *ragged-opsom* (*Ra^{Op}*) C57BL/6J mice were obtained from the Jackson Laboratory (Bar Harbor) and interbred to generate *Ra^{Op}/Ra^{Op}* embryos, *Ra^{Op}/+* adults and wild-type littermates for analysis. *Sox18*-null embryos were genotyped as previously described¹⁴. Backcrossing of *Sox18* null mice from the original mixed background (129/B6) was accomplished by 11 generations of backcrossing onto a pure B6 background. *Prox1*-GFP transgenic mouse embryos were produced by pronuclear injection into CBA/C57BL/6 F1 zygotes and screened by PCR. Procedures involving animals and their care conformed to institutional guidelines (University of Queensland Animal Ethics Committee).

Immunofluorescence. Immunofluorescence was performed as previously described²⁶. Skin and ear samples were dissected from 16.5 d.p.c. transgenic embryos and adult mice respectively, and fixed overnight in 4% PFA. Primary antibodies in blocking solution were added and incubated for 24 h at 4 °C. Samples were washed for 6 h in washing solution (1% BSA, 1% DMSO in PBSTx) and incubated for 16 h with secondary antibodies in blocking solution. Confluent H5V cells were fixed in 4% PFA for 15 min at room temperature. Cells were permeabilized with PBSTx (0.3%) for 5 min at 4 °C and then blocked in PBS/5%BSA for 30 min at room temperature.

Quantitation of lymphatic vasculature. After immunofluorescence using anti-LYVE-1 or anti-PECAM-1 antibody, apical photographs of mouse ears were taken at 10× magnification. The numbers of lymphatic vessel branching points, lymphatic vessel loops, and 'blind ended' lymphatic sacs were determined, in a blinded fashion, and averaged for an entire image, as described previously²⁷.

Antibodies. Antibodies were used at the following dilutions: rabbit polyclonal anti-mouse *Prox1* (Covance), 1:2,500; hamster anti-mouse podoplanin (RDI), 1:500; rabbit polyclonal anti-mouse *Sox18* (Supplementary Fig. 9) raised against peptides PLAPEDCALRAFRA and RPYAPELARDPSFC (a.a. 228–241 and 240–254 respectively of mouse *Sox18*), biotinylated as described²⁵, 1:1,000; rabbit polyclonal anti-mouse LYVE-1 (Fitzgerald Industries), 1:1,000; rat anti-mouse CD31 (BD Pharmingen), 1:200; mouse monoclonal anti- α -tubulin (clone B-5-1-2, Sigma) 1:2,000; mouse monoclonal NG2 antibody (Chemicon) 1:200; mouse monoclonal anti-MYC Tag (9B11, Cell Signalling), 1:2,000; rabbit polyclonal anti-GFP (Molecular Probes), 1:1,000; rabbit polyclonal anti-Lbx2 (Millipore) 1:1,000; rat monoclonal anti-endoglin (BD Pharmingen), 1:200; goat polyclonal anti-neuropilin-1 (R&D Systems) 1:50; hamster polyclonal anti-ICAM-1 (Millipore) 1:200; mouse monoclonal anti-N-cadherin, 1:200 (kindly provided by A. Yap); rat monoclonal anti-CD44 (BD Pharmingen) 1:200; rat monoclonal anti-mouse VE-cadherin (BD Pharmingen) 1:300; and rat monoclonal anti-CD34 as previously described²⁸. Secondary antibodies anti-rat IgG Alexa 594, anti-rabbit IgG Alexa 488, anti-hamster IgG Alexa 488, anti-biotin IgG Alexa 594, and anti-biotin IgG Alexa 488 (Molecular Probes) were used at a dilution of 1:200, and anti-rabbit IgG HRP conjugated (Zymed Laboratories) used at a dilution of 1:2,000.

ES cell culture. ES cells were differentiated to endothelial cells as previously described^{16,17}. ES cells were mildly trypsinized and suspended in Iscove's modified Dulbecco medium with 15% FBS, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin, 450 μ M monothio glycerol, 10 μ g ml⁻¹ insulin, 50 ng ml⁻¹ human recombinant VEGF-A₁₆₅ (Peprotech Inc.), 2 U ml⁻¹ human recombinant erythropoietin (Cilag AG), and 100 ng ml⁻¹ human bFGF (Genzyme). Cells were seeded in bacteriological Petri dishes and cultured for 3, 7 or 11 days at 37 °C with 5% CO₂ and 95% relative humidity.

Quantitative PCR analysis. cDNA was amplified in triplicate in a reaction volume of 15 μ l using TaqMan Gene Expression Assay (Applied Biosystems) and an ABI/Prism 7900 HT thermocycler, using a pre-PCR step of 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 60 s at 60 °C. Preparations of RNA template without reverse transcriptase were used as negative controls. *C_t* values were normalized to *Gapdh* expression²⁹.

Electrophoretic mobility shift assay. EMSA was carried out as previously described¹⁹. Two double-stranded oligonucleotides were constructed, containing

either of the SOX binding sites (underlined) and surrounding sequences from the *Prox1* promoter at -1,135 bp (*SoxA*) 5'-GGTTCCCCCGCCCCAGACAAT-GCTAGTTTGCATACAAAG-3' and at -813 bp (*SoxB*) 5'-ACAGTCCATCTCC-CTTTCTAACAAATTGAAGTCACCAATG-3'.

Plasmids. The mouse *Prox1* promoter fragment encompassing nucleotides -3,952 to -1 from the ATG was generated by PCR. Hot-start PCR was performed (annealing temperature at 58 °C). The resulting fragment cloned into the *KpnI* and *XhoI* sites of the pGL2Basic luciferase reporter vector (Promega). The 4 kb *Prox1* promoter fragment was subcloned into pEGFP-N1 vector (Clontech, BD Bioscience) which had the CMV promoter removed. The integrity of the reporter construct was verified by sequencing. *SoxA* and *SoxB* mutant clones were generated by site directed mutagenesis (QuickChange, Stratagene) using mutated primers (5'-CCCCCGCCCCAGATGATGCTAGTTTGC-3') at position -1,135 or a different mutated primer (5'-GGCAACAGTCCATCTCCCTTTCTAATGATTGAAGTCACC-3') at position -813.

Cell culture and transfection experiments. Mouse mesenteric lymph node endothelial cells (mlEnd) were created as previously described²⁰ and obtained from J. Gamble. Cells were transfected using a combination of Lipofectamine 2000 (Invitrogen) and CombiMag (Chemicell) for 20 min (according to manufacturer's instructions) on a magnetic plate at room temperature. Cells were harvested 24 h after transfection, and luciferase assays were carried out as previously described¹². H5V murine endothelial cells were cultured as previously described¹⁸.

Lentiviral vector preparation. To generate HIV3-PGK vector, the *Pgk* promoter from pRetro/SUPER (ref. 30) was ligated into the lentiviral plasmid pRRLsin.PPTs.hCMV.GFPwpre (ref. 31) in place of the CMV promoter. *Sox18* and *Sox18^{RaOp}* cDNAs were subcloned in HIV3-PGK and lentiviral vectors produced as described³². Lentiviral and packaging plasmids were kindly donated by L. Naldini.

Chromatin immunoprecipitation. A total of 1 × 10⁷ mlEnd cells were cross-linked with 1% formaldehyde for 10 min at room temperature. Chromatin immunoprecipitation was performed as previously described²⁶. Primers 5'-GGCAAGGGTGTTCCTTTT-3' and 5'-GGCACAGACATGCACTGTCC-3' used in a 22 cycle PCR reactions (annealing temperature 50 °C).

X-gal staining. X-gal staining was performed as previously described³³.

Western blotting. TM3 cells were transfected with *Sox7*, *Sox17* or *Sox18* expression constructs. Proteins were extracted in 2× sample buffer (125 mM Tris, pH 6.8; 4% SDS; 20% glycerol and 5% β -mercaptoethanol). Primary antibody was used overnight at 4 °C. HRP-conjugated secondary antibody followed by chemiluminescence detection with Super Signal West Pico Chemiluminescence reagent (Pierce) were used to reveal protein expression.

26. Wilhelm, D. *et al.* SOX9 regulates prostaglandin D synthase gene transcription *in vivo* to ensure testis development. *J. Biol. Chem.* **282**, 10553–10560 (2007).
27. Shayan, R. *et al.* A system for quantifying the patterning of the lymphatic vasculature. *Growth Factors* **25**, 417–425 (2007).
28. Garlanda, C. *et al.* Characterization of MEC 14.7, a new monoclonal antibody recognizing mouse CD34: A useful reagent for identifying and characterizing blood vessels and hematopoietic precursors. *Eur. J. Cell Biol.* **73**, 368–377 (1997).
29. Spagnuolo, R. *et al.* Gas1 is induced by VE-cadherin and vascular endothelial growth factor and inhibits endothelial cell apoptosis. *Blood* **103**, 3005–3012 (2004).
30. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550–553 (2002).
31. Follenzi, A. *et al.* Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nature Genet.* **25**, 217–222 (2000).
32. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72**, 8463–8471 (1998).
33. Leung, K. K. *et al.* Different cis-regulatory DNA elements mediate developmental stage- and tissue-specific expression of the human COL2A1 gene in transgenic mice. *J. Cell Biol.* **141**, 1291–1300 (1998).

LETTERS

Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin

Emma Byres^{1*}, Adrienne W. Paton^{2*}, James C. Paton², Jonas C. Löfling³, David F. Smith⁴, Matthew C. J. Wilce¹, Ursula M. Talbot², Damien C. Chong², Hai Yu⁵, Shengshu Huang⁵, Xi Chen⁵, Nissi M. Varki³, Ajit Varki³, Jamie Rossjohn¹ & Travis Beddoe¹

AB₅ toxins comprise an A subunit that corrupts essential eukaryotic cell functions, and pentameric B subunits that direct target-cell uptake after binding surface glycans. Subtilase cytotoxin (SubAB) is an AB₅ toxin secreted by Shiga toxigenic *Escherichia coli* (STEC)¹, which causes serious gastrointestinal disease in humans². SubAB causes haemolytic uraemic syndrome-like pathology in mice³ through SubA-mediated cleavage of BiP/GRP78, an essential endoplasmic reticulum chaperone⁴. Here we show that SubB has a strong preference for glycans terminating in the sialic acid *N*-glycolylneuraminic acid (Neu5Gc), a monosaccharide not synthesized in humans. Structures of SubB–Neu5Gc complexes revealed the basis for this specificity, and mutagenesis of key SubB residues abrogated *in vitro* glycan recognition, cell binding and cytotoxicity. SubAB specificity for Neu5Gc was confirmed using mouse tissues with a human-like deficiency of Neu5Gc and human cell lines fed with Neu5Gc. Despite lack of Neu5Gc biosynthesis in humans, assimilation of dietary Neu5Gc creates high-affinity receptors on human gut epithelia and kidney vasculature. This, and the lack of Neu5Gc-containing body fluid competitors in humans, confers susceptibility to the gastrointestinal and systemic toxicities of SubAB. Ironically, foods rich in Neu5Gc are the most common source of STEC contamination. Thus a bacterial toxin's receptor is generated by metabolic incorporation of an exogenous factor derived from food.

The B subunits of AB₅ toxins typically recognize cognate glycan receptors displayed on cell-surface glycoconjugates^{5,6}. Receptor specificity is critical for the pathogenic process, as it determines host susceptibility, tissue tropism, and the nature and spectrum of the resultant pathology. Accordingly, we sought to understand the receptor specificity of SubAB. Glycan array analysis showed that Oregon-Green-labelled SubAB (OG–SubAB) had a high degree of binding specificity for glycans terminating with α 2-3-linked residues of the non-human sialic acid Neu5Gc (Table 1). Much weaker binding was seen with those glycans that terminated in α (2-3)-linked *N*-acetylneuraminic acid (Neu5Ac), which differs by one hydroxyl group from Neu5Gc (Fig. 1a). Table 1 is a list of glycans selected from the microarray analysis of SubAB toxin and a mutant derivative SubAB_{A12} (discussed later). This list represents the glycans on the array to which native SubAB has the highest apparent affinity and corresponding Neu5Ac derivatives, asialo- and sulphated-derivatives. Of all the glycans on the array, Neu5Gc α 2-3Gal β 1-4-GlcNAc β (#260) bound SubAB best. The binding of SubAB to this glycan is reduced 20-fold if the Neu5Gc is changed to Neu5Ac (#237); over 30-fold if the Neu5Gc linkage is changed from α 2-3 to α 2-6

(#263); and 100-fold if the sialic acid is removed (#152). The high binding of SubAB to structures #258, #260 and #261 indicates that it has a high affinity for terminal α 2-3-linked Neu5Gc, with little discrimination for the penultimate moiety. Surface plasmon resonance analysis (Supplementary Fig. 2a) showed an approximately tenfold higher SubAB binding response to Neu5Gc α 2-3Lac β than to Neu5Ac- α (2-3)-Lac- β . Competitive inhibition studies (Supplementary Fig. 2b) indicated that Neu5Gc α 2-3Lac β has an inhibition constant (K_i) of 2 mM, which is in the range reported for other monovalent sialic-acid–protein interactions^{7,8}. This high specificity of SubAB for Neu5Gc-terminating glycans is unique among bacterial toxins.

Next we determined the structure of the *apo*-form of the SubB pentamer (Fig. 1 and Supplementary Table 1). As expected the SubB protomer adopted the common oligonucleotide/oligosaccharide-binding fold⁹, typical of other AB₅ toxins. The SubB structure most resembled the S2, S3 and S5 subunits of pertussis toxin (PTX; Supplementary Table 2), where the S2/3 subunits of PTX contained a shallow binding site for sialylated glycoproteins¹⁰; SubB also contained a similar shallow pocket lined by similar residues. In all such previously studied AB₅ toxins, the sialic acid in question is Neu5Ac, a common sialic acid found in humans and other animals. We then determined the structure of SubB in complex with free Neu5Gc (Supplementary Table 1). Neu5Gc bound to SubB unambiguously (Fig. 1c), whereas identical experiments using Neu5Ac failed to show any binding. The shallow binding pocket of SubB bound Neu5Gc in the chair conformation (Fig. 1d). The Neu5Gc lay in the α -anomeric configuration, even though most free Neu5Gc (more than 90%) in solution is in the β -anomeric form; SubB is therefore highly selective for the small fraction of α -anomeric Neu5Gc present in solution. Neu5Gc is coordinated mainly by multiple polar interactions with SubB (Supplementary Table 3). The predominant van der Waals interaction arose from Phe 11, which swivelled by 90° when Neu5Gc was bound, stacking parallel with the sugar ring and forming a cap over the binding site. In addition, the Neu5Gc was sequestered by many direct and water-mediated hydrogen bonds, including interactions with the side chains of Asp 8, Ser 12, Glu 36 and Tyr 78. Neu5Gc differs from Neu5Ac by the addition of a hydroxyl on the methyl group of the *N*-acetyl moiety of Neu5Ac. This extra hydroxyl present in Neu5Gc made crucial interactions with SubB: namely, the extra hydroxyl interacts with Tyr 78^{OH} and hydrogen bonds with the main chain of Met 10. These key interactions could not occur with Neu5Ac, thus explaining the marked preference for Neu5Gc.

¹Protein Crystallography Unit and ARC Centre of Excellence for Structural and Functional Microbial Genomics, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia. ²School of Molecular and Biomedical Science, University of Adelaide, South Australia 5005, Australia. ³Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California 92093-0687, USA. ⁴Protein-Carbohydrate Interaction Core H, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ⁵Department of Chemistry, University of California, Davis, California 95616, USA.

*These authors contributed equally to this work.

Table 1 | Glycan array analysis of native SubAB and B subunit mutant SubAB_{A12}

Glycan number	Structure	SubAB		SubAB _{A12}	
		Mean relative fluorescence units	Coefficient of variation (%)	Mean relative fluorescence units	Coefficient of variation (%)
260	Neu5Gcα2-3Galβ1-4GlcNAcβ-Sp0*	37,606	4	590	10
261	Neu5Gcα2-3Galβ1-4Glcβ-Sp0	25,197	16	1,998	30
264	Neu5Gcα-Sp8*	25,114	19	172	87
258	Neu5Gcα2-3Galβ1-3GlcNAcβ-Sp0	24,209	12	327	54
15	α-Neu5Ac-Sp11	8,836	30	1,242	46
31	[3OSO3]Galβ1-3(Fucα1-4)GlcNAcβ-Sp8	3,094	10	1,375	41
33	[3OSO3]Galβ1-3GlcNAcβ-Sp8	3,006	15	3,808	45
231	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp8	2,401	8	1,587	72
46	Neu5Acα2-3[6OSO3]Galβ1-4GlcNAcβ-Sp8	2,254	16	2,997	11
36	[3OSO3]Galβ1-4GlcNAcβ-Sp0	1,923	22	3,426	4
34	[3OSO3]Galβ1-4(Fucα1-3)GlcNAcβ-Sp8	1,878	36	2,398	32
237	Neu5Acα2-3Galβ1-4GlcNAcβ-Sp8	1,862	31	4,247	8
240	Neu5Acα2-3Galβ1-4Glcβ-Sp8	1,799	10	1,127	25
243	Neu5Acα2-6GalNAcβ1-4GlcNAcβ-Sp0	1,730	18	1,871	67
117	Galβ1-3(Fucα1-4)GlcNAc-Sp0	1,549	42	1,395	55
68	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ-Sp8	1,385	30	1,204	53
225	Neu5Acα2-3Galβ1-3GlcNAcβ-Sp0	1,315	75	1,563	16
257	Neu5Gcα2-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp0	1,244	36	2,250	14
259	Neu5Gcα2-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	1,163	37	1,719	11
263	Neu5Gcα2-6Galβ1-4GlcNAcβ-Sp0	1,154	34	859	16
245	Neu5Acα2-6Galβ1-4GlcNAcβ-Sp0	1,048	34	1,989	38
262	Neu5Gcα2-6GalNAcα-Sp0	1,003	10	125	23
37	[3OSO3]Galβ1-4GlcNAcβ-Sp8	989	21	1,362	63
217	Neu5Acα2-3Galβ1-3(Fucα1-4)GlcNAcβ-Sp8	922	14	335	42
10	GalNAcα-Sp8	857	39	938	43
230	Neu5Acα2-3Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	798	26	344	24
134	Galβ1-3GlcNAcβ-Sp8	678	121	197	33
246	Neu5Acα2-6Galβ1-4GlcNAcβ-Sp8	658	30	755	26
154	Galβ1-4Glcβ-Sp0	644	41	789	48
118	Galβ1-3(Fucα1-4)GlcNAc-Sp8	604	31	411	59
242	Neu5Acα2-6GalNAcα-Sp8	564	16	434	24
133	Galβ1-3GlcNAcβ-Sp0	539	25	604	24
239	Neu5Acα2-3Galβ1-4Glcβ-Sp0	484	46	318	23
49	9-O-AcNeu5Acα2-6Galβ1-4GlcNAcβ-Sp8	478	42	204	59
4	Neu5Gcβ2-6Galβ1-4GlcNAc-Sp8	471	24	176	45
92	GalNAcβ1-4GlcNAcβ-Sp0	419	15	235	24
152	Galβ1-4GlcNAcβ-Sp0	374	40	755	27
67	Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ-Sp0	326	52	273	103
155	Galβ1-4Glcβ-Sp8	299	32	164	73
153	Galβ1-4GlcNAcβ-Sp8	298	20	123	44

Data are presented for a selection of 40 of the 320 glycans present on the array. Complete data sets are available at: www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp. Data are for quadruplicate array spots.

* Sp0 and Sp8 designate CH₂CH₂NH₂ and CH₂CH₂CH₂NH₂ linkers, respectively.

Given the glycan array data, we also established how a trisaccharide that terminated in Neu5Gc (Neu5Gcα2-3Galβ1-3GlcNAc; #258 in the array) bound SubB (Fig. 1e). The mode of binding of the Neu5Gc moiety in the monosaccharide and the trisaccharide complex was identical. The remaining two sugar moieties present in the trisaccharide extended to solvent, but were nevertheless able to contact SubB. The small number of additional interactions between SubB and the tertiary sugar in the trisaccharide (Fig. 1e) is consistent with the Neu5Gc moiety driving the specificity and affinity for the interaction with SubB. Of note, the sequence Neu5Gcα2-3Galβ1-3GlcNAc is very common as a terminating structure of *N*-glycans in non-human animals.

Despite the commonality of oligonucleotide/oligosaccharide-binding fold, neither the receptor specificity nor the location of the receptor-binding site is conserved throughout the AB₅ toxin family. For Shiga and cholera toxins, whose receptors are glycolipids, the deep receptor-binding pockets are located on the membrane face of the toxin^{11–13}. Like SubB¹⁴, PTX binds sialylated glycoproteins; moreover, the PTX S2/3 sialic-acid binding site is also shallow and in the same location, halfway down the sides of the pentamer (Fig. 1f–h). Examination of the SubB–Neu5Gc structure superposed onto the structure of PTX in complex with the disaccharide, Neu5Acα2-3Gal, reveals that both toxins bind sialic acid in the same orientation, with similar residues interacting with the sialic-acid head group (Fig. 1d, f). However, the interactions between Tyr 78 and

Asp 8 of SubB with the extra hydroxyl group of Neu5Gc have no such equivalents in PTX. At the same position as Tyr 78 in SubB is the small, non-polar Val 167 side chain in PTX S2/3, and there are no residues that overlay in the PTX S2/3 region equivalent to Asp 8 in SubB. The absence of these side chains in PTX S2/3 provides a basis for understanding why PTX, unlike SubB, is not specific for Neu5Gc.

SubB residues that we considered critical for Neu5Gc binding were then mutated to determine their role in biologically relevant toxin–receptor interactions. The most critical residue was Ser 12, mutation of which abolishes interactions with the C1 carboxylate group of sialic acid; this reduced cytotoxicity for Vero cells by 99.98% (Supplementary Table 4). In addition, binding of labelled mutated toxin (OG–SubAB_{A12}) to Vero cells could not be detected by fluorescence microscopy, confirming an effect on receptor recognition (Fig. 2a). Further glycan array analysis using OG–SubAB_{A12} indicated markedly reduced binding to the immobilized sialylated glycans (Table 1). The elimination of the hydroxyl of Tyr 78 in the SubAB_{F78} mutant holotoxin also reduced Vero cell cytotoxicity by 96.9% (Supplementary Table 4), which highlights the importance of interactions with the hydroxyl group unique to Neu5Gc-containing glycans. The trace residual cytotoxicity of the SubAB_{F78} mutant is probably attributable to SubB still being able to bind weakly to Neu5Ac-containing glycans, because Tyr 78 would not be required for binding Neu5Ac. Labelled SubAB_{F78} also exhibited reduced binding to Vero cells (Fig. 2a). The Gln36Ala mutation resulted in

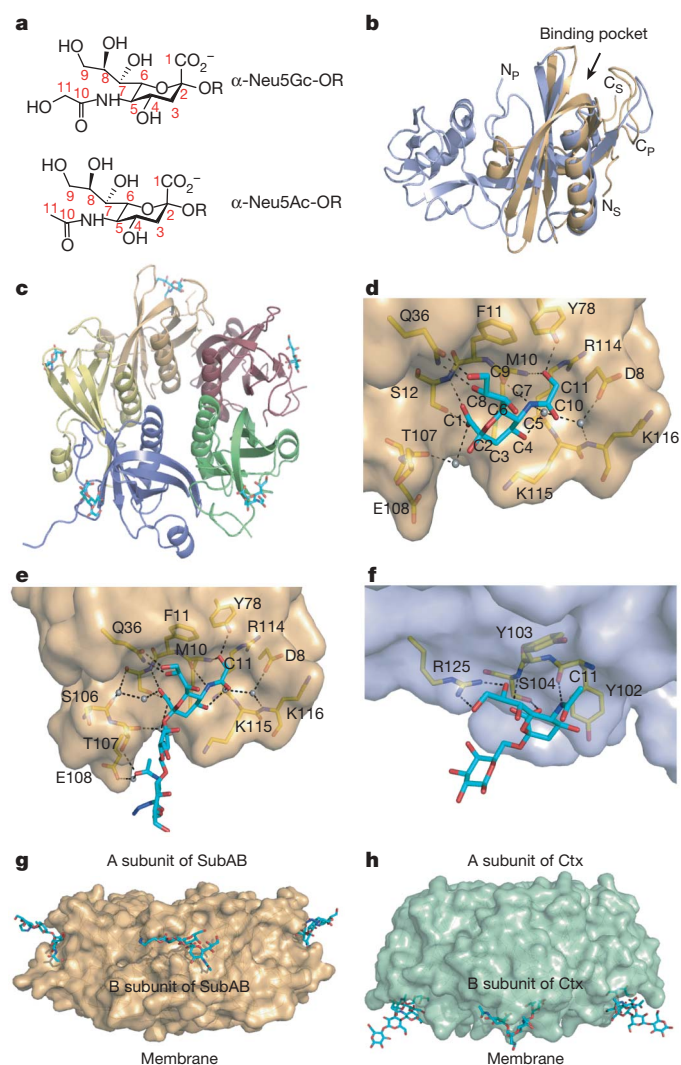


Figure 1 | Structural analysis of SubB-sialic acid interactions and comparison with other AB₅ toxins. **a**, Structures of Neu5Gc and Neu5Ac, showing additional O at C11 of the former. **b**, SubB protomer (orange) superimposed upon PTX S2 (blue). Amino (N) and carboxy (C) termini of SubB and PTX S2 are designated by subscripts S and P, respectively. **c**, Representation of the pentameric SubB-Neu5Gc structure, with each protomer colour-coded. Cyan sticks represent the sugar, blue sticks represent nitrogen atoms, and red sticks represent oxygen atoms. **d**, Neu5Gc in sialic acid receptor-binding site of SubB. The extra hydroxyl of Neu5Gc interacts with Tyr78^{OH}, and hydrogen bonds with the main chain of Met 10. **e**, Trisaccharide Neu5Gcα2-3Galβ1-3GlcNAcβProN₃ binding to SubB. ProN3 refers to the linker used in the synthesis. **f**, Neu5Acα2-3Gal binding site of PTX S2/3, which shares similarity to the SubB binding site, namely: Ser 12 in SubB, Ser 104 in PTX S2/3; Gln 36 in SubB, Arg 125 in PTX S2/3; Phe 11 in SubB, Tyr 103 in PTX S2/3. PTX does not have the equivalent of Tyr 78 and Asp 8. **g**, Side-on view of SubB:Neu5Gcα2-3Galβ1GlcNAcβProN₃. **h**, Side-on view of CtxB:GM₁. In **d-h**, cyan sticks represent ligands, dark blue sticks represent nitrogen atoms, and red sticks oxygen atoms. Yellow sticks represent key residues in the protein backbone. Black dotted lines represent hydrogen bonds.

reduction of cytotoxicity by 88% (Supplementary Table 4), which is consistent with the importance of the interactions with the C8 and C9 hydroxyl groups present on both Neu5Gc and Neu5Ac.

As further evidence of the biological significance of the high specificity of SubAB for Neu5Gc-terminating glycans, we showed that OG-SubAB bound to kidney tissue from wild-type mice, but not cytidine monophosphate (CMP)-*N*-acetylneuraminic acid hydroxylase (*Cmah*)-null mice¹⁵ that had a human-like genetic defect in the

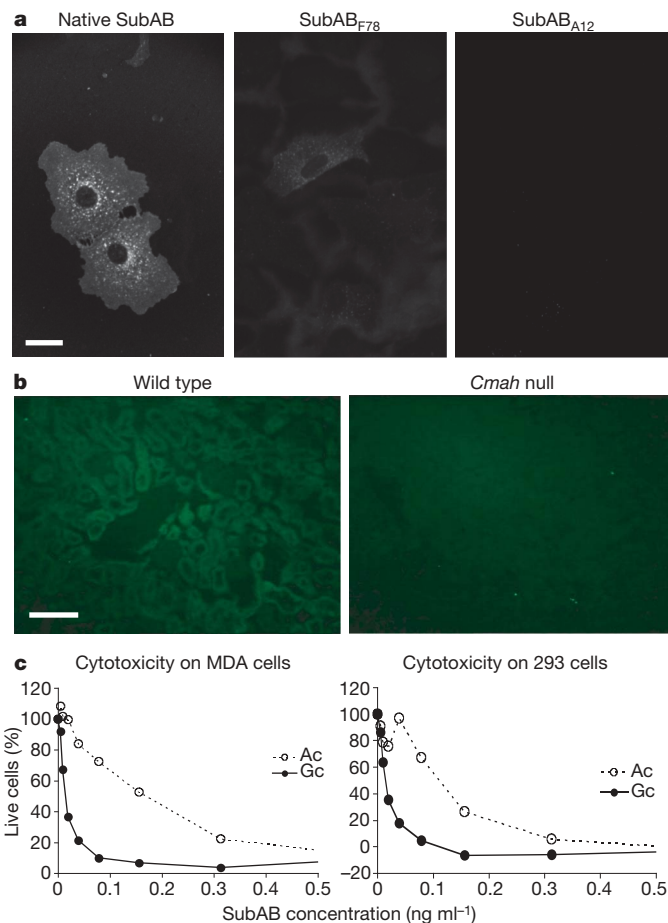


Figure 2 | Fluorescence microscopy and Neu5Gc-dependent cytotoxicity. **a**, Binding of SubAB mutants to Vero cells. Vero cells growing on coverslips were incubated with 1 $\mu\text{g ml}^{-1}$ OG-SubAB, OG-SubAB_{A12} or Texas Red-SubAB_{F78} for 60 min at 37 °C (scale bar, 25 μm). **b**, Binding of SubAB to kidney sections from wild-type and *Cmah*-null mice. Frozen kidney sections were incubated with 1 $\mu\text{g ml}^{-1}$ OG-SubAB for 60 min at 37 °C, washed, and examined by epifluorescence microscopy (scale bar, 100 μm). **c**, Neu5Gc-dependent cytotoxicity of SubAB for MDA-MB-231 cells or 293 cells after growth in medium supplemented with 3 mM Neu5Gc (Gc) or Neu5Ac (Ac) was determined as described in the Methods. Data shown are from a single experiment. Analysis of data pooled from triplicate experiments yielded CD₅₀ values of 3.92 ± 1.58 pg (mean \pm standard error) and 13.17 ± 2.46 pg for MDA cells fed Neu5Gc or Neu5Ac ($P = 0.034$; two-tailed *t*-test). For 293 cells fed Neu5Gc or Neu5Ac the CD₅₀ values are 1.33 ± 0.30 pg and 11.58 ± 3.30 pg, respectively ($P = 0.036$). The mean increase in SubAB susceptibility (\pm standard error) for cells fed Neu5Gc versus Neu5Ac was 5.56 ± 2.71 -fold for MDA cells and 8.35 ± 0.88 -fold for 293 cells.

ability to convert CMP-Neu5Ac to CMP-Neu5Gc (Fig. 2b). Humans lack this enzyme owing to a mutation in the *Cmah* gene that occurred after evolutionary separation of the hominin lineage from the great apes¹⁶, suggesting the possibility of human genetic resistance to the toxin. However, human cells can metabolically assimilate Neu5Gc present in tissue culture media and incorporate it into cell-surface glycans¹⁷. We therefore manipulated the levels of Neu5Gc on the surface of human cell lines (MDA-MB-231 breast cancer cells and 293 embryonic kidney cells), which had been adapted for growth in human serum, resulting in undetectable amounts of Neu5Gc. These cells were then grown for 3 days in medium supplemented with 3 mM Neu5Gc or 3 mM Neu5Ac. In the former case, this increased Neu5Gc content to 50–75% of total sialic acid in membrane preparations (result not shown). Feeding MDA and 293 cells with Neu5Gc rather than Neu5Ac significantly increased their susceptibility to SubAB (50% cytotoxic dose (CD₅₀) values decreased 5.56- and

8.35-fold, respectively), rendering them as sensitive as Vero cells (Fig. 2c).

We have previously shown that some normal human tissues contain small quantities of Neu5Gc¹⁷. No tissues in *Cmah*-null mice express Neu5Gc, as determined by a highly specific polyclonal chicken antibody, and by mass spectrometry¹⁵. Thus, human Neu5Gc must be derived from dietary sources, a mechanism confirmed by previous studies of human volunteers¹⁷. We therefore

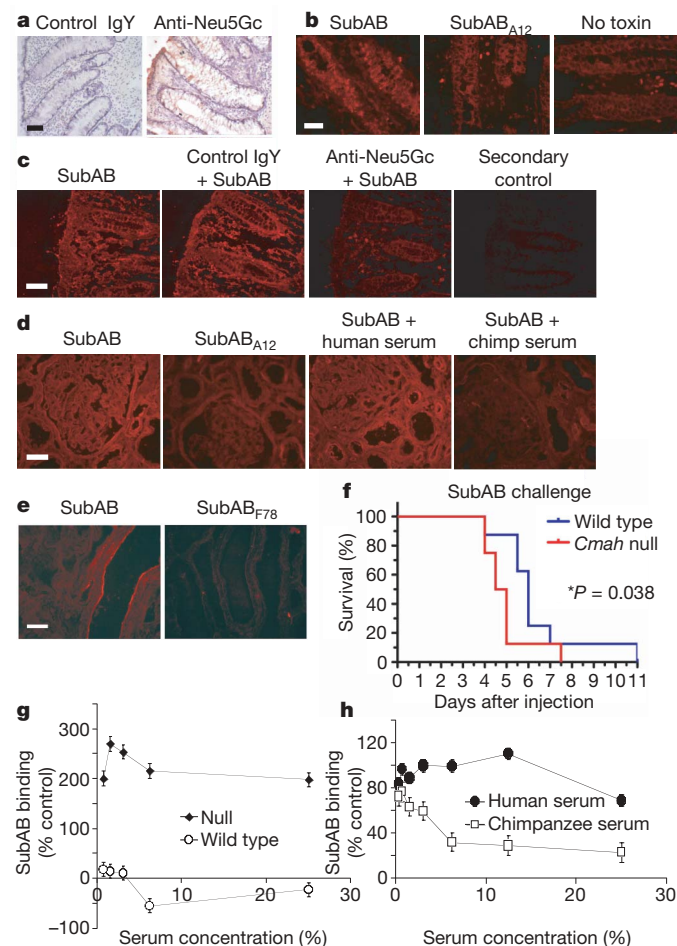


Figure 3 | Neu5Gc-dependent binding of SubAB to human tissues and toxicity of SubAB in wild-type and *Cmah*-null mice. **a**, Frozen sections of human colon were stained with chicken anti-Neu5Gc or control IgY at $5 \mu\text{g ml}^{-1}$, followed by anti-chicken IgY–HRP conjugate, and examined by immunohistochemistry¹⁷ (scale bar, $100 \mu\text{m}$). **b**, Similar human colon sections were overlaid with or without $1 \mu\text{g ml}^{-1}$ SubAB or SubAB_{A12} and bound toxin was detected using rabbit anti-SubA and Cy3-labelled goat anti-rabbit IgG and examined by epifluorescence microscopy (see Methods) (scale bar, $50 \mu\text{m}$). **c**, Human colon sections were overlaid first with anti-Neu5Gc or control IgY at $5 \mu\text{g ml}^{-1}$, followed by $1 \mu\text{g ml}^{-1}$ SubAB, and bound toxin was detected as for **b**. Background control sections received only rabbit anti-SubAB, followed by Cy3-labelled anti-rabbit IgG (scale bar, $100 \mu\text{m}$). **d**, **e**, Human kidney sections were overlaid $1 \mu\text{g ml}^{-1}$ SubAB, SubAB_{A12} or SubAB_{F78}, and in the presence or absence of 10% human or chimpanzee serum, as indicated. Bound toxin was detected as for **b** (scale bar, $50 \mu\text{m}$). **f**, Wild-type and *Cmah*-null mice ($n = 8$ each) were injected intraperitoneally with 200 ng g^{-1} purified SubAB in a total of $100 \mu\text{l}$ PBS, and survival time was recorded. Kaplan–Meier survival curves were plotted and statistical analysis (Wilcoxon–Gehan test) was performed in GraphPad Prism. **g**, **h**, Inhibition of SubAB binding to immobilized Neu5Gc- $\alpha(2-3)$ -Lac-HSA by wild-type versus *Cmah*-null mouse serum (**f**), or human versus chimpanzee serum (**g**) was assayed by enzyme-linked immunosorbent assay (ELISA) as described in the Methods. Data are expressed as the percentage of a control with no serum and are the mean \pm standard error of triplicate wells and representative of three independent experiments.

studied human colon sections by using an IgY antibody with absolute specificity for Neu5Gc (unpublished improvements over ref. 15), and noted that Neu5Gc was present on the epithelial surfaces and in the crypts (Fig. 3a). Similar regions of human colon sections also bound native SubAB, but binding of SubAB_{A12} was markedly diminished (Fig. 3b). Importantly, pre-incubation with anti-Neu5Gc IgY, but not control IgY, substantially blocked binding of SubAB to human colon sections (Fig. 3c), demonstrating competition for the same epitope. Specific binding of SubAB to human kidney sections, particularly the glomerular endothelium, was also observed, but this was not seen for SubAB_{A12} (Fig. 3d). Human kidney vasculature also shows staining with the anti-Neu5Gc antibody (data not shown). Furthermore, binding to kidney tissue was not seen with SubAB_{F78}, which is defective only in Neu5Gc-specific receptor interactions (Fig. 3e). Thus, binding of SubAB to human colon and kidney tissue is Neu5Gc dependent.

We also examined whether the inability of *Cmah*-null mice to express Neu5Gc affected *in vivo* susceptibility to injected SubAB. Surprisingly, the null mice had a slightly shorter median survival time (5 versus 6 days, $P = 0.038$; Fig. 3f). We reasoned that because normal mouse serum contains high levels of Neu5Gc-containing glycoproteins, these would compete with receptors on the surface of tissues for SubAB, thereby providing partial protection against toxicity. In contrast, human and *Cmah*-null mouse serum would not have any such protective activity. Thus, in the null mouse, lack of protective serum glycoproteins could counteract any benefit derived from lack of expression of Neu5Gc on cell-surface glycoconjugates. Indeed, wild-type but not *Cmah*-null mouse serum competitively inhibited binding of SubAB to immobilized Neu5Gc-glycan (Fig. 3g). Similar results were obtained for chimpanzee serum (which, like wild-type mouse serum, is rich in Neu5Gc-glycoproteins) versus human serum (Fig. 3h). Furthermore, chimpanzee serum, but not human serum, inhibited binding of SubAB to human kidney sections (Fig. 3d). Given that chimpanzee serum glycoproteins otherwise share near identity with their human counterparts, the effects can be attributed to the presence or absence of Neu5Gc competitors¹⁸.

This study presents the first example of a bacterial toxin showing a marked preference for Neu5Gc-containing glycans and is potentially significant in the context of host susceptibility to toxin-mediated disease. Indeed, glycans terminating in Neu5Gc $\alpha(2-3)\text{Gal}\beta(1-3)\text{GlcNAc}$ sequences are widely expressed on the cells of many mammals including livestock, suggesting an evolutionary reason for the emergence of this selective binding preference. Neu5Gc is not produced by bacteria or plants, is low or absent in poultry and fish, but abundant in red meats (lamb, pork and beef) and in bovine milk. However, humans are the known exception among mammals, and Neu5Ac predominates because of the *Cmah* mutation¹⁶. How then could the toxin mediate disease in humans? The answer probably lies in the diet, as human ingestion of red meats and milk products results in Neu5Gc incorporation into human tissues¹⁷. We detected significant levels of Neu5Gc on human colonic epithelium, and Neu5Gc-specific binding of SubAB at this site, as well as in the kidney vasculature, the major target organ for haemolytic–uraemic syndrome caused by STEC (Fig. 3). This Neu5Gc can only have originated from the diet, as there is no known alternative pathway for Neu5Gc biosynthesis¹⁵. Ironically, red meat and dairy products (the richest dietary sources of Neu5Gc) are the very foods that are most commonly contaminated with SubAB-producing STEC². Thus, through regular dietary intake of red meats and milk, humans may pre-sensitize their tissues to a key virulence factor of a major pathogen that occurs sporadically in the same foods. Furthermore, because of the absence of protective Neu5Gc-bearing glycoproteins in their serum and other body fluids, humans who have consumed foods with high Neu5Gc content may actually be hyper-susceptible to the toxin, as illustrated in Supplementary Fig. 1.

METHODS SUMMARY

Purification and labelling of SubAB. SubAB holotoxin, and derivatives with B-subunit mutations, were purified by Ni-NTA chromatography and labelled with Oregon Green, as described previously^{19,20}.

Structure determination. SubB was crystallized, and the Se-Met-labelled SubB structure was solved using the multiple anomalous dispersion technique. Crystal soaks and difference Fourier analysis was used to locate the Neu5Gc-containing binding sites. Further details are provided in the Methods.

Glycan array analysis. Binding of OG-labelled SubAB and mutant derivatives thereof to immobilized glycans was investigated using a printed array of 320 glycan targets on version 3.0 of the glycan microarray of the Consortium for Functional Glycomics Core H (<http://www.functionalglycomics.org/static/consortium/resources/resourcecoreh8.shtml>). Further details are provided in the Methods.

DNA methods. Routine DNA manipulations were performed essentially as described previously¹. DNA sequencing used dye-terminator chemistry and an ABI 3700 sequencer.

Cell culture and cytotoxicity assays. SubAB and mutant derivatives were assayed for cytotoxicity on Vero cells as previously described¹. Cytotoxicity was also assayed on MDA-MB-231 (human breast cancer) and 293 (human embryonic kidney) cells, which had been adapted for growth in human serum to eliminate presence of Neu5Gc. Cells were then cultured for 3 days in 96-well plates in medium (RPMI and DMEM, respectively) supplemented with 3 mM Neu5Gc or Neu5Ac, and then exposed to serial dilutions of SubAB, as described for Vero cells¹. Cell viability was assessed after incubation for a further 3 days by staining with 10% Alamar blue (Serotec), according to the manufacturer's instructions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 June; accepted 15 September 2008.

Published online 29 October 2008.

- Paton, A. W., Srimanote, P., Talbot, U. M., Wang, H. & Paton, J. C. A new family of potent AB₅ cytotoxins produced by Shiga toxigenic *Escherichia coli*. *J. Exp. Med.* **200**, 35–46 (2004).
- Paton, J. C. & Paton, A. W. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Clin. Microbiol. Rev.* **11**, 450–479 (1998).
- Wang, H., Paton, J. C. & Paton, A. W. Pathologic changes in mice induced by subtilase cytotoxin, a potent new *Escherichia coli* AB₅ toxin that targets the endoplasmic reticulum. *J. Infect. Dis.* **196**, 1093–1101 (2007).
- Paton, A. W. & Beddoe, T. Thorpe, C.M., Whisstock, J.C., Wilce, M.C.J., Rossjohn, J., Talbot, U.M. and Paton J.C. AB₅ subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. *Nature* **443**, 548–552 (2006).
- Sandvig, K. & van Deurs, B. Membrane traffic exploited by protein toxins. *Annu. Rev. Cell Dev. Biol.* **18**, 1–24 (2002).
- Lencer, W. I. & Tsai, B. The intracellular voyage of cholera toxin: going retro. *Trends Biochem. Sci.* **28**, 639–645 (2003).
- Crocker, P. R., Paulson, J. C. & Varki, A. Siglecs and their roles in the immune system. *Nature Rev. Immunol.* **7**, 55–66 (2007).
- Neu, U., Woellner, K., Gauglitz, G. & Stehle, T. Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc. Natl Acad. Sci. USA* **105**, 5219–5224 (2008).
- Murzin, A. G. OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861–867 (1993).
- Stein, P. E. *et al.* Structure of a pertussis toxin-sugar complex as a model for receptor binding. *Nature Struct. Biol.* **1**, 591–596 (1994).
- Merritt, E. A., Sixma, T. K., Kalk, K. H., van Zanten, B. A. & Hol, W. G. Galactose-binding site in *Escherichia coli* heat-labile enterotoxin (LT) and cholera toxin (CT). *Mol. Microbiol.* **13**, 745–753 (1994).
- Merritt, E. A. *et al.* Structural studies of receptor binding by cholera toxin mutants. *Protein Sci.* **6**, 1516–1528 (1997).
- Ling, H., Bast, D. & Brunton, J. L. & R. e. a. d. R. J. Structure of the Shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37**, 1777–1788 (1998).
- Yahiro, K. *et al.* Identification and characterization of receptors for vacuolating activity of subtilase cytotoxin. *Mol. Microbiol.* **62**, 480–490 (2006).
- Hedlund, M. *et al.* N-glycolylneuraminic acid deficiency in mice: implications for human biology and evolution. *Mol. Cell. Biol.* **27**, 4340–4346 (2007).
- Varki, A. Multiple changes in sialic acid biology during human evolution. *Glycoconjugate J.* doi:10.1007/s10719-008-9183-z (7 September 2008).
- Tangvoranuntakul, P. *et al.* Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl Acad. Sci. USA* **100**, 12045–12050 (2003).
- Gagneux, P. *et al.* Proteomic comparison of human and great ape blood plasma reveals conserved glycosylation and differences in thyroid hormone metabolism. *Am. J. Phys. Anthropol.* **115**, 99–109 (2001).
- Talbot, U. M., Paton, J. C. & Paton, A. W. Protective immunization of mice with an active-site mutant of subtilase cytotoxin of Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **73**, 4432–4436 (2005).
- Chong, D. C., Paton, J. C., Thorpe, C. M. & Paton, A. W. Clathrin-dependent trafficking of subtilase cytotoxin, a novel AB₅ toxin that targets the ER chaperone BiP. *Cell. Microbiol.* **10**, 795–806 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff at the General Medicine and Cancer Institutes Collaborative Access Team Advanced Photon Source, Chicago, for assistance with data collection. This research was supported by a Program Grant from the National Health and Medical Research Council of Australia (NHMRC; to A.W.P. and J.C.P.), an NHMRC Project Grant (to T.B. and A.W.P.), a grant from the National Institute of General Medical Sciences to the Consortium for Functional Glycomics, RO1 grants from the National Institutes of Health (to A.W.P., J.C.P., J.R., X.C. and A.V.) and from the ARC Centre of Excellence in Structural and Functional Microbial Genomics (to J.R.). J.R. is supported by an Australian Research Council Federation Fellowship; T.B. by an NHMRC Career Development Award; J.C.P. by an NHMRC Australia Fellowship. We also thank C. J. Gregg for assistance with *in vivo* experiments and with collection of data, and L. Wiggleton for technical assistance with tissue sectioning and staining.

Author Contributions E.B. and A.W.P. contributed equally. E.B., T.B. and M.C.J.W. crystallized SubB, solved the structure and contributed to manuscript preparation. A.W.P. constructed mutants and contributed to design and interpretation of experiments, project management and writing of the manuscript. J.C.P. and J.R. contributed to design and interpretation of experiments, project management and writing of the manuscript. D.C.C. and U.M.T. performed experiments. D.F.S. performed and interpreted glycan array experiments. J.C.L., N.M.V. and A.V. designed, performed and interpreted experiments relating to Neu5Gc on cells and tissues and to cytotoxicity *in vivo*, and contributed to manuscript preparation. H.Y., S.H. and X.C. synthesized oligosaccharides. A.V., T.B. and J.R. are joint senior and corresponding authors.

Author Information The coordinates and structure factors for the SubB structures are deposited in Protein Data Bank under accession numbers 3DWA, 3DWP and 3DWQ. Raw glycan array data are available at www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.V. (a1varki@ucsd.edu), J.R. (jamie.rossjohn@med.monash.edu.au) or T.B. (travis.beddoe@med.monash.edu.au).

METHODS

Glycan array analysis. Labelled proteins were diluted to 0.1 mg ml^{-1} in Tris-buffered saline (TBS: 20 mM Tris, 150 mM NaCl, 2 mM CaCl_2 , 2 mM MgCl_2 , pH 7.4) containing 1% BSA and 0.05% Tween-20, and an aliquot (70 μl) was applied to separate microarray slides and incubated under a coverslip for 60 min at room temperature. Coverslips were then removed and slides were washed by dipping four times in successive washes of TBS containing 0.05% Tween-20, TBS and deionized water. Slides were then spun for approximately 15 s to dry and immediately scanned in a PerkinElmer ProScanArray MicroArray Scanner using an excitation wavelength of 488 nm and ImaGene software (BioDiscovery) to quantify fluorescence. The data are reported as average relative fluorescence units of four of six replicates (after removal of the highest and lowest values) for each glycan represented on the array.

Purification and crystallization of SubB. Purification of SubB was performed as detailed previously for the holotoxin⁴ with the following modifications. SubB was eluted from an Ni-NTA column using wash buffer containing 500 mM imidazole and further purified by gel filtration using an S200 16/60 column (GE Healthcare) pre-equilibrated in 20 mM Tris HCl, 150 mM NaCl, 2 mM EDTA, pH 7.0. Fractions containing SubB were pooled and purity was assessed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry (data not shown). SubB was concentrated to 2 mg ml^{-1} and 40 mM 3-(decyldimethylammonio) propanesulphonate was added before storing at 4°C . Crystals of SubB grew in drops containing 1 μl protein (2 mg ml^{-1}) and 1 μl reservoir solution, with the reservoir solution consisting of 500 μl of 16% (w/v) PEG 3350, 100 mM sodium cacodylate, pH 6.2, and 200 mM ammonium fluoride. Selenomethionine-substituted SubB was purified and crystallized under the same conditions.

X-ray data collection, structure determination and refinement. X-ray data were collected for SeMet-labelled SubB at the GMCA-CAT beamline at the Advanced Photon Source, Chicago. Data were collected from a single crystal at three wavelengths: inflection point, peak and high energy remote. Native SubB crystals were soaked by the addition of 1 mM of Neu5Gc, Neu5Ac or the trisaccharide Neu5Gc α 2-3Gal β 1-3GlcNAc β ProN₃ for 1 h at 20°C . Datasets from the soaked crystals were collected in-house using an R-axis IV⁺⁺ detector. X-ray data were autoindexed and processed with MOSFLM²¹, SCALA²² and the CCP4 suite of programs²³.

The *apo*-form of SubB was phased by the single-wavelength anomalous dispersion (SAD) method using SeMet inflection point data and the PHENIX Autosol suite²⁴. An initial model of 460 out of 630 residues was built by RESOLVE²⁵, and this was used as input for ARP/wARP²⁶. The model was completed by rounds of manual building in COOT²⁷ followed by refinement in REFMAC5 (ref. 28) with fivefold non-crystallographic symmetry maintained until the final stages of refinement. Waters and PEG 400 molecules were added in COOT to give a final SubB model consisting of a homopentamer of 588 residues (with each B subunit between 115 and 120 residues). The *apo*-SubB structure was used, through PHASER²⁹, to solve the sugar-bound structures, with residues rebuilt and refined as described for the *apo* structure. Ligands were constructed using the PRODRG2 server³⁰. Clear density was observed for the various sugars in the SubB binding sites and these were further confirmed by constructing $F_o - F_c$ omit maps. Data collection and refinement statistics are shown in Supplementary Table 1. The stereochemistry and overall quality of each of the structures was confirmed by CCP4 programs.

Site-directed mutagenesis of *subB*. Derivatives of SubAB with either S12A, Q36A or Y78F substitutions in the B subunit were constructed using overlap extension PCR mutagenesis. For the S12A mutant, this involved high-fidelity PCR amplification (Expand High Fidelity PCR kit; Roche Molecular Diagnostics) of *subAB*-positive O113:H21 STEC DNA using primer pairs pETsubAF/SubBA12R and SubBA12F/pETsubBR (Supplementary Table 5). This generated two fragments with the necessary mutation in the S12 codon of *subB* incorporated into the overlapping region by the SubBA12R and SubBA12F primers. The two separate PCR products were purified, mixed together and the complete *subAB* region re-amplified using primer pair pETsubAF/pETsubAR. These primers incorporate BamHI and XhoI restriction sites, enabling direct cloning into pET-23(+) (Novagen), followed by transformation into *E. coli* BL21(DE3) (Novagen). The other mutations were constructed in an analogous

manner, using first-round primer pairs pETsubAF/SubBA36R and SubBA36F/pETsubBR for Q36A, and pETsubAF/SubBF78R and SubBF78F/pETsubBR for Y78F. Mutations were confirmed by sequencing, and expression of intact SubAB protein was confirmed by SDS–PAGE and western blotting. Holotoxins carrying confirmed B-subunit mutations were purified as for native toxin.

Tissue immunohistochemical studies. Human colon or kidney sections were overlaid with or without $1 \mu\text{g ml}^{-1}$ SubAB, SubAB_{A12} or SubAB_{F78}, with or without 10% human or chimpanzee serum, as indicated, for 1 h at room temperature. Slides were then washed and fixed in 10% buffered formalin for 15 min and washed again. The sections were overlaid with rabbit anti-SubA at 1:5,000 in 1% BSA/PBS and incubated for 1 h at room temperature, washed again and overlaid with Cy3-labelled goat anti-rabbit IgG (Jackson ImmunoResearch) for 1 h at room temperature. Slides were washed, mounted using aqueous mounting medium and viewed using epifluorescence with a Zeiss Axiophot microscope with appropriate excitation and barrier filters. Digital photomicrographs were taken using a Sony CCD (charge-coupled device) camera and National Institutes of Health image software and photo panels constructed using Adobe Photoshop and Illustrator.

SubAB binding inhibition ELISA. Inhibition of SubAB binding to immobilized Neu5Gc α 2-3LacHSA by wild-type versus *Cmah*-null mouse serum, or human versus chimpanzee serum, was assayed as follows. All reagents were diluted and blocking was done in 1% fish gelatin. Costar 96-well ELISA plates were coated with saturating amounts of Neu5Gc α 2-3LacHSA at 4°C overnight, and then blocked at 4°C overnight. Serum was added at the indicated concentrations, followed by 15 ng ml^{-1} SubAB. After 2 h incubation at room temperature, the plates were washed four times in PBS. Rabbit anti-SubAB serum diluted 1:5,000 was added. The plates were incubated for 2 h at room temperature, followed by another washing step, and then horseradish peroxidase conjugated to goat anti-rabbit IgG, 1:20,000, for 2 h at room temperature. After washing, 3,3',5,5'-tetramethylbenzidine (Sigma) was added to the plates; the reaction was stopped after 15 min by 2 M sulphuric acid and absorbance was read at 450 nm.

Surface plasmon resonance. Experiments were conducted at 25°C on a Biacore 3000 instrument using HBS buffer (10 mM HEPES–HCl, pH 7.4, 800 mM NaCl and 0.005% surfactant P20; supplied by the manufacturer). Approximately 200 resonance units of Neu5Ac α 2-3Lac β -biotin and Neu5Gc α 2-3Lac β -biotin was immobilized onto streptavidin-coupled sensor chips (Biacore). SubAB was passed over all flow cells at $20 \mu\text{l min}^{-1}$ for 1 min. The final response was calculated by subtracting the response of the control surface from the glycan surface. For inhibition studies, SubAB (100 nM) was incubated with increasing concentrations of Neu5Gc α 2-3Lac β -biotin (3 μM to 4 mM) for 1 h at room temperature before being passed over the chip at $5 \mu\text{l min}^{-1}$ for 3 min. After each injection, the surface was regenerated with three injections of 10 mM glycine, pH 2.0. The experiments were performed in duplicate. The amount of SubAB bound at equilibrium was used to generate the inhibition curve that was analysed by nonlinear regression using PRISM software (version 3.0).

- Leslie, A. G. W. in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography* 26 (SERC, 1992).
- Evans, P. R. in *Proc. CCP4 Study Weekend on Recent Advances in Phasing* 97–102 (CCLRC, 1997).
- Collaborative Computational Project. Number 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
- Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**, 38–44 (2003).
- Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. Warp: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy models. *Acta Crystallogr. D* **30**, 551–554 (1997).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Murshodov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
- McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61**, 458–464 (2005).
- Schuettkopf, A. W. & van Aalten, D. M. F. PRODRG – a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. D* **60**, 1355–1363 (2004).

Crystal structure of the ZP-N domain of ZP3 reveals the core fold of animal egg coats

Magnus Monné^{1*}, Ling Han^{1*}, Thomas Schwend¹, Sofia Burendahl¹ & Luca Jovine¹

Species-specific recognition between the egg extracellular matrix (zona pellucida) and sperm is the first, crucial step of mammalian fertilization¹. Zona pellucida filament components ZP3 and ZP2 act as sperm receptors, and mice lacking either of the corresponding genes produce oocytes without a zona pellucida and are completely infertile². Like their counterparts in the vitelline envelope of non-mammalian eggs and many other secreted eukaryotic proteins, zona pellucida subunits polymerize using a 'zona pellucida (ZP) domain' module^{3–5}, whose conserved amino-terminal part (ZP-N) was suggested to constitute a domain of its own⁶. No atomic structure has been reported for ZP domain proteins, and there is no structural information on any conserved vertebrate protein that is essential for fertilization and directly involved in egg–sperm binding. Here we describe the 2.3 ångström (Å) resolution structure of the ZP-N fragment of mouse primary sperm receptor ZP3. The ZP-N fold defines a new immunoglobulin superfamily subtype with a β -sheet extension characterized by an E' strand and an invariant tyrosine residue implicated in polymerization. The structure strongly supports the presence of ZP-N repeats within the N-terminal region of ZP2 and other vertebrate zona pellucida/vitelline envelope proteins, with implications for overall egg coat architecture, the post-fertilization block to polyspermy and speciation. Moreover, it provides an important framework for understanding human diseases caused by mutations in ZP domain proteins and developing new methods of non-hormonal contraception.

A three-dimensional reconstruction of ZP domain protein endoglin has been recently reported at 25 Å resolution⁷. However, crystallographic studies of this family of proteins have long been hampered by their intrinsic polymerization activity and extensive post-translational modifications, including several intramolecular and intermolecular disulphides, highly heterogeneous N/O-glycosylation, and proteolytic cleavage of precursor forms^{5,8}. Expression of isolated ZP-Ns has not resulted in soluble protein, but fusion to N-terminal *Escherichia coli* maltose binding protein (MBP, also known as MalE) allowed partially soluble expression of the ~11 kDa ZP-N fragment of mouse ZP3 (residues Val 42–Arg 143) in bacteria, engineered to form disulphide bonds in the cytoplasm⁶. The carboxy-terminally polyhistidine-tagged protein—which is 71% identical to its human counterpart but does not have glycosylation sites—contained the invariant Cys₁–Cys₄ and Cys₂–Cys₃ disulphide bonds (in which the subscripts number the conserved Cys residues) that are the hallmark of the ZP domain^{3,5,6,8,9}. However, it could not be crystallized because it formed filamentous aggregates *in vitro*⁶.

ZP domain proteins have a highly mosaic architecture, and very few residues often separate ZP-N sequences from the variable domains that immediately precede them^{3,5}. We reasoned that minimization of the linker between MBP and ZP-N, a strategy used to

favour crystallization of challenging proteins¹⁰, might shield surfaces causing protein aggregation without affecting the folding of ZP-N. By selectively eluting monomeric molecules from the affinity column used during purification, we obtained milligram amounts of a tri-Ala linker construct (MBP-AAA-ZP-N-6His) that was considerably more soluble and less aggregated than the original fusion (Supplementary Fig. 1a–f). This protein crystallized readily (Supplementary Fig. 1g–i) and its structure was solved in three crystal forms by molecular replacement with MBP (form I), molecular replacement with MBP-AAA-ZP-N (form II) and combined molecular replacement/Zn-single-wavelength anomalous diffraction (SAD) phasing (form III) (Supplementary Table 1 and Supplementary Figs 2 and 3).

ZP-N has approximate dimensions of 52×29×28 Å. Eight β -strands, accounting for 57% of the sequence, form an immunoglobulin-like antiparallel β -sandwich that could not be predicted from the sequence because of its unusual features and topology (Fig. 1). Whereas β -sheet N1 consists of classical immunoglobulin-fold β -strands A–B–E–D¹¹, strands F and G of β -sheet N2 are much longer than strand C such that they extend beyond the β -sandwich and, together with an extra E' strand, form a platform (E'–F–G extension) at the bottom of the molecule. The E' strand, which is preceded by a single-turn α -helix ($\alpha_{EE'}$) and forms a β -bulge with Ser 112 in strand F, extends in the same direction as strand C. A highly structured 11-residue loop (BC) and a single-turn 3_{10} helix (3_{10BC}) connect β -strands B and C at the top of ZP-N, where a second long loop (FG) is located that is poorly defined in the electron density.

The ZP-N hydrophobic core consists of residues belonging to β -strands B, E, C and the upper half of F and G; β -strand A residues Val 42 and Val 44 also contribute to the core in ZP3 homologues (Fig. 1c). The structure is stabilized by the two invariant disulphides of the ZP domain that clamp both sides of the β -sandwich and, except for part of Cys₁46, are completely buried at the bottom of the hydrophobic core (Fig. 2 and Supplementary Fig. 2b). Structure-based alignments indicate that, in other proteins, the ZP-N fold is reinforced by additional disulphides in β -sheet N2; furthermore, they suggest that the ordered loop BC can mediate interactions between ZP-N and preceding domains, because it often contains one extra Cys residue that could form an interdomain disulphide (Supplementary Table 2).

The finding that ZP-N adopts an immunoglobulin-like fold demonstrates that it is a domain of its own. At the same time, ZP-N has only low overall structural similarity to other immunoglobulin-like domains, and even its closest match (D3 domain of ICAM-1 (ref. 12); secondary-structure matching (SSM) Q-score 0.29) is considerably different in topology, hydrophobic core and disulphide linkages. Together with the absence of an E' strand in previously described immunoglobulin-like domains, this suggests that ZP-N belongs to a new immunoglobulin superfamily subtype.

¹Karolinska Institutet, Department of Biosciences and Nutrition, Hälsovägen 7, SE-141 57 Huddinge, Sweden.

*These authors contributed equally to this work.

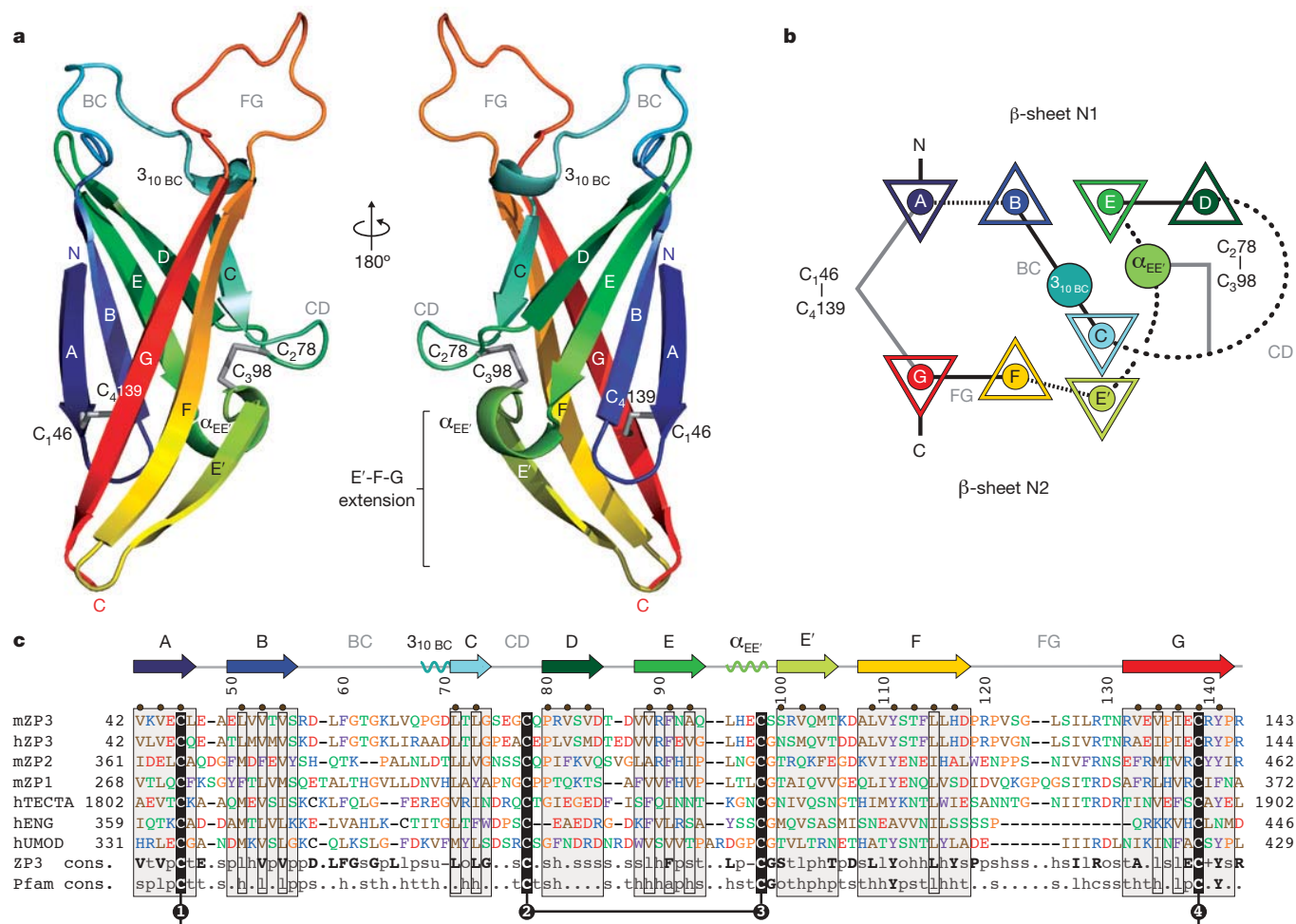


Figure 1 | Overall structure of the ZP-N domain of ZP3. a, Cartoon representation rainbow-coloured from blue (N terminus) to red (C terminus), with conserved disulphides shown as grey sticks. **b**, Topology, with strands as triangles, and helices as circles. Connections between secondary structure elements at the top, middle or bottom of the structure are represented by straight continuous, rounded dashed and straight dashed lines, respectively. **c**, Structure-based alignment of ZP-N and mouse

(m) and human (h) zona pellucida proteins, as well as non-egg coat ZP domain proteins α -tectorin (TECTA), endoglin (ENG) and uromodulin (UMOD). Consensus sequences (cons.) for ZP3 homologues and Pfam ZP domain family seed sequences are also shown (for colour-coding and consensus keys, see Methods). Brown circles mark residues in which the side chains lie on the inner side of β -sheets; open and closed boxes indicate conserved hydrophobic core and Cys residues, respectively.

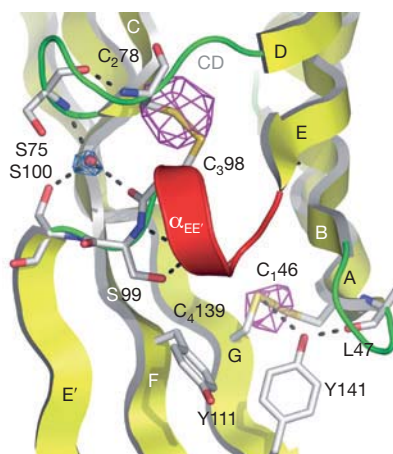


Figure 2 | Key contacts involving conserved ZP-N domain residues. View of the bottom of the hydrophobic core and E'-F-G extension of ZP3 ZP-N, with important residues shown as sticks and hydrogen bonds indicated by dashed lines. Violet mesh is a 4.7 Å resolution phased anomalous difference map contoured at 3.5 σ , calculated using diffraction data collected at 8 keV to confirm disulphide bond assignments. Blue mesh is a 2.3 Å resolution composite annealed omit map, contoured at 1.0 σ and cropped around a water molecule that mediates interaction between Cys₃₉₈, Ser 75 and Ser 100.

Whereas the C-terminal fragment of the ZP domain (ZP-C) of ZP3 has been implicated in sperm binding, ZP-N is involved in zona pellucida filament formation^{2,6}. Incorporation of mouse ZP3 into the *Xenopus* vitelline envelope suggests that sequence and structural features important for egg coat assembly are maintained throughout evolution¹³. Two highly conserved regions of ZP-N (Fig. 3a, ovals) expose hydrophobic residues and do not overlap with sites that can be glycosylated or are protease-accessible (Fig. 3b and Supplementary Table 3). Nonspecific juxtaposition of β -strands A, G and the hydrophobic C terminus of MBP (Supplementary Fig. 4) explains why shortening of the MBP/ZP-N linker reduces protein aggregation (Supplementary Fig. 1f). The second region, essentially corresponding to the ZP-N-specific E'-F-G extension, forms intermolecular contacts through antiparallel pairing of symmetry-related E' β -strands in crystal form III (Fig. 3c) and mediates stacking of molecular layers in the I222 and P1 crystals (Supplementary Fig. 3). Moreover, it is the part of the structure most strongly predicted to be involved in protein-protein interactions (Supplementary Fig. 5). The extension contains Tyr 111 and Tyr 141 that are involved in an aromatic-aromatic interaction and are the only ZP-N residues, apart from the Cys, to be nearly invariant (Figs 1c and 2). Both tyrosines are missing in ZP domain protein endoglin that does not form polymers¹⁴ (Fig. 1c), and human deafness-associated mutation of the residue corresponding to Tyr 111 in inner ear protein α -tectorin (Tyr1870Cys) disrupts formation of

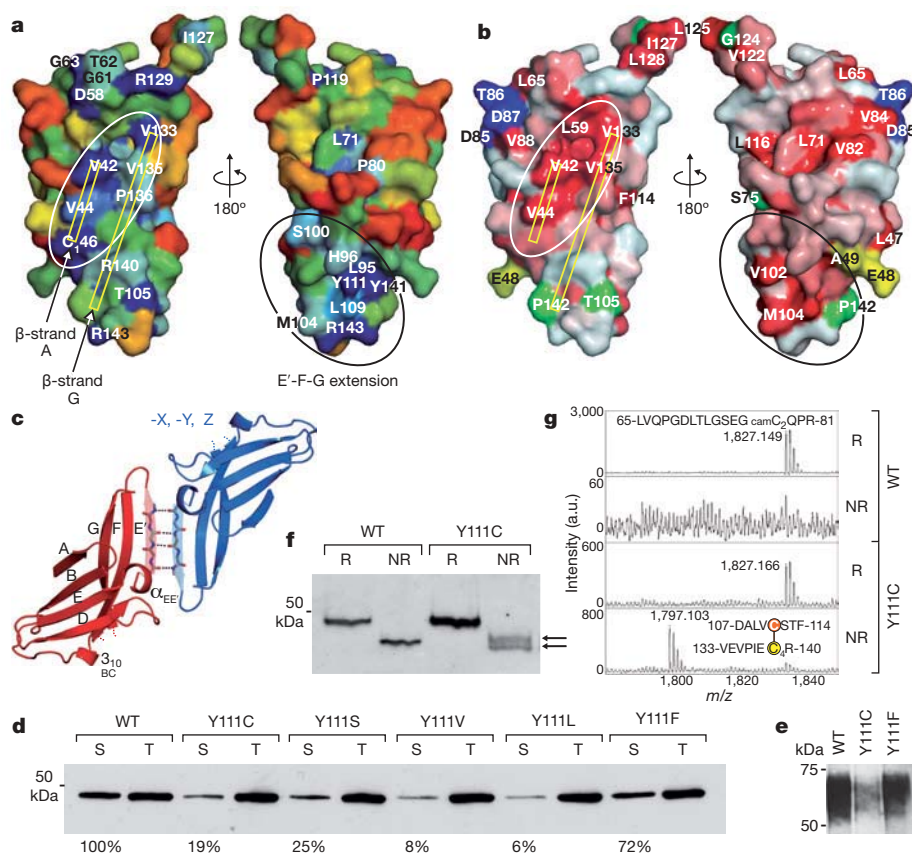


Figure 3 | ZP-N domain interaction sites. **a, b**, Surface representation of ZP3 ZP-N, with residues coloured from red to blue by increasing conservation (**a**) or from white to red by increasing hydrophobicity (**b**). N-glycosylation, protease cleavage and RGD sites are green, yellow and blue, respectively (see Supplementary Table 3). Views are from the MBP interface (left) and opposite side (right). **c**, Intermolecular backbone interactions between ZP-N-specific E' β -strands connect N2 β -sheets in crystal form III. **d**, Immunoblot analysis of soluble (S) and total (T) expression of wild-type (WT) MBP-AAA-ZP-N-6His and Tyr 111 mutants with monoclonal anti-5His antibody. Percentages indicate the amount of soluble protein relative to the wild type. **e**, Immunoblot analysis with

polyclonal anti-ZP3 antibody of conditioned medium from Chinese hamster ovary (CHO)-K1 cells expressing full-length ZP3 constructs. **f**, Immunoblot analysis of soluble wild-type and Tyr111Cys MBP-AAA-ZP-N-6His with anti-5His antibody in reducing (R) and non-reducing (NR) conditions. For comparison, tenfold more extract was loaded for the mutant than the wild type. **g**, Matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) analysis of trypsin/chymotrypsin-digested samples from **f** detects an abnormal Cys 111–Cys₄139 disulphide in non-reduced Tyr111Cys mutant protein only. The intense peak at m/z 1827, corresponding to carbamidomethylated (cam) ZP3 peptide Leu 65–Arg 81, confirms efficient disulphide reduction after dithiothreitol treatment.

striated-sheet matrix filaments in the sulcal region of the tectorial membrane¹⁵. As also observed on substitution with other non-aromatic residues, mutation of Tyr 111 to Cys reduces solubility of MBP-AAA-ZP-N-6His (Fig. 3d), and a corresponding decrease in secretion is seen in full-length ZP3 Tyr111Cys expressed in mammalian cells (Fig. 3e). On the other hand, Tyr111Phe mutants of MBP-AAA-ZP-N-6His and full-length ZP3 are ~70% soluble (Fig. 3d) and almost as efficiently secreted (Fig. 3e) as the respective wild-type proteins, in agreement with the observation that the Tyr is naturally substituted by Phe in a very small number of ZP domain sequences. Impairment of protein secretion is a general feature of most other characterized ZP-N domain mutations, which cause disease by haploinsufficiency owing to effects on folding that can be explained on the basis of the structure (Supplementary Fig. 6 and Supplementary Table 4). Unlike these mutations, the Tyr-to-Cys substitution must however also directly affect filament assembly, because the tectorial membrane is not disrupted in normal-hearing animals heterozygous for a targeted deletion in α -tectorin (*Tecta*^{+/ Δ ENT})¹⁵. Tyr 111 is adjacent to Cys₄139 in the structure (Fig. 2), suggesting that the mutation could interfere with disulphide bond formation. Analysis of MBP-AAA-ZP-N-6His Tyr111Cys shows that this consists of two disulphide isoforms (Fig. 3f, arrows), one of which contains an abnormal Cys111–Cys₄139 bridge (Fig. 3g). In molecular dynamics simulations both isoforms are as stable as the wild type when folded and show only

local differences where the Tyr side chain is absent (Supplementary Fig. 7). In light of studies on the effect of the mutation on protein assembly into the zona pellucida⁴, these data collectively indicate that, although only a small fraction of α -tectorin Tyr1870Cys folds and is secreted, this acts as a dominant negative by incorporating into nascent tectorial membrane filaments and preventing their elongation due to structural alterations around the conserved Tyr. Thus, crystallographic, mutational and biochemical evidence demonstrates a crucial role for the ZP-N-specific E'–F–G extension in polymerization.

Our structure has implications for animal egg coat architecture by strongly supporting the recent hypothesis that zona pellucida/vitelline envelope proteins contain additional N-terminal copies of ZP-N¹⁶. Using structure-based distant homologue recognition, we identified three isolated ZP-N domains in the N-terminal region of secondary sperm receptor ZP2 and built a three-dimensional model of its entire ZP-N1–N4 array (Fig. 4a, b). The model is structurally sound because it buries conserved hydrophobic residues and, consistent with the absence of free Cys residues in ZP2 (ref. 8), involves most Cys in canonical Cys₁–Cys₄, Cys₂–Cys₃ ZP-N bridges that agree with available disulphide bond assignments^{8,17,18}. The remaining two Cys are positioned so that they can form an extra disulphide within ZP-N3 (Fig. 4c). By exposing sites known to be glycosylated, targeted by proteases, or recognized by monoclonal antibodies, the model is also entirely consistent with a large amount of other experimental data available on ZP2

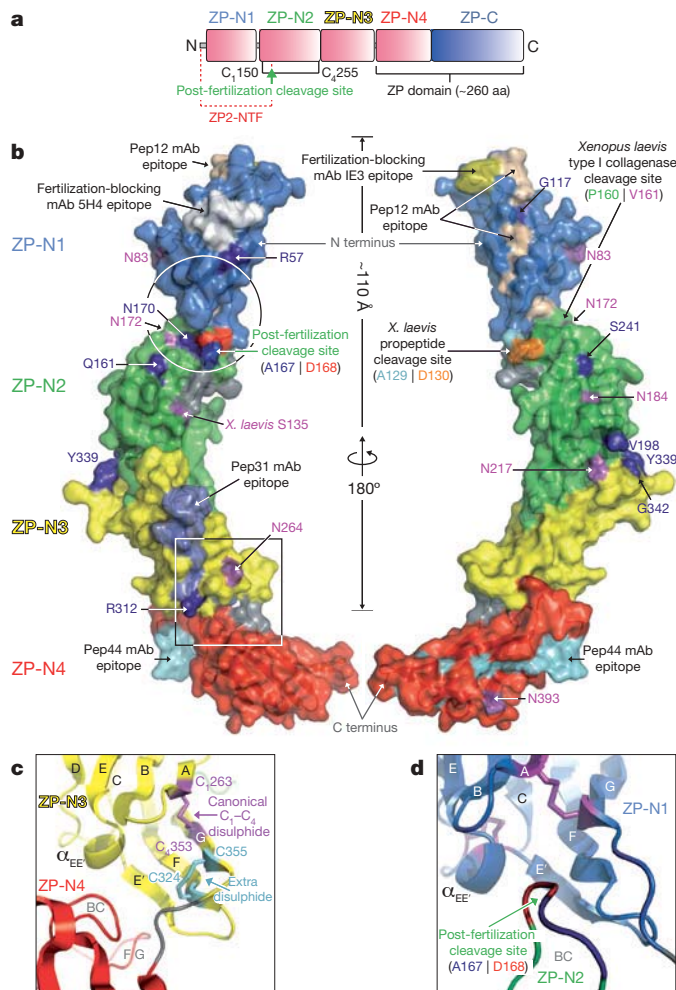


Figure 4 | Model of the ZP-N domain repeat region of ZP2. **a**, **b**, Domain architecture of ZP2 (**a**) and three-dimensional surface model of its ZP-N1–N4 region (**b**), with relevant features indicated (see Supplementary Table 5). aa, amino acids. Glycosylation and positively selected sites are dark blue and violet, respectively. Unless otherwise specified, the residue numbers refer to mouse ZP2 sequence. mAb, monoclonal antibody. **c**, Detail of the region boxed in **b**, highlighting a further disulphide bond between non-canonical Cys residues of ZP-N3. **d**, View of the region circled in **b**, suggesting that post-fertilization cleavage of ZP-N2 BC loop could affect the position of ZP-N1 relative to the rest of ZP2.

(Fig. 4b and Supplementary Table 5). Furthermore, owing to short or absent interdomain linkers, ZP-N1–N4 assumes an extended rod-like structure, the ~110 Å span of which matches the size of protrusions extending from isolated zona pellucida filaments imaged by electron microscopy¹⁹. This agrees with the observation that the C-terminal ZP domain of ZP2 is responsible for its polymerization into filaments, from which the N-terminal fragment (ZP-N1–N3) could protrude⁴; accordingly, the latter was shown to bind acrosome-reacted sperm²⁰, providing a target for contraceptive vaccines^{21,22}.

After gamete fusion, site-specific proteolysis of ZP2 produces an N-terminal fragment (ZP2-NTF) that remains attached to the rest of the protein via a single disulphide bond involving its fifth Cys^{17,18,23}. Cleavage of ZP2 triggers hardening of the zona pellucida as part of the block to polyspermy^{17,23,24} and was recently proposed to regulate mammalian sperm binding to the zona pellucida by modulating the supramolecular structure of the matrix²⁴. On the basis of our model, the ZP2 cleavage site lies within the long BC loop of ZP-N2, whose canonical Cys₁ corresponds to the fifth Cys of ZP2-NTF; thus, after fertilization the latter remains linked to the rest of ZP2 through the Cys₁–Cys₄ disulphide of the second repeat (Fig. 4a, b, d). Proteolytic processing could alter the orientation of ZP-N1

relative to ZP-N2 because the substrate loop lies at the interface between the two repeats, in proximity to the ZP-N1/ZP-N2 linker (Fig. 4d). At the same time, cleavage of the only loop connecting the two β -sheets at the top of ZP-N2 would make the underlying region available for interaction. Considering that other extracellular adhesion molecules often establish cell–cell contacts through the formation of adhesive zippers^{12,25}, this combination of events could trigger contacts between ZP-N repeats protruding from different zona pellucida filaments. This would result in the observed post-fertilization three-dimensional compaction of filaments in the inner layer of the zona pellucida²⁶, generating a physical barrier to sperm penetration as well as mask binding sites for sperm on ZP-N1–N3.

Extension of the structure-based analysis that we applied to ZP2 to other zona pellucida/vitelline envelope proteins indicates that the ZP-N fold described in this study accounts for >50% of the structure of all egg coats from frog to human. However, we could not identify extra copies of ZP-N in ZP domain proteins not involved in fertilization, or in vitelline envelope components from fish whose only function is structural. Thus ZP-N repeats are only found in egg coat subunits of vertebrate classes where the zona pellucida/vitelline envelope is directly responsible for species-restricted gamete interaction (Supplementary Fig. 8). ZP2 and ZP3 are among the 10% most divergent proteins in mammals²⁷ and, on the basis of our model, 11 out of 12 sites suggested to be positively selected in the sequence of mature ZP2 (refs 28 and 29) are located within ZP-N repeats 1–3. Most of these residues are predicted to be exposed and could potentially be available for interaction with sperm (Fig. 4b and Supplementary Table 5). Taken together, these observations raise an unexpected parallel with invertebrate gamete recognition proteins, where rapid evolution of repeated domains is implicated in speciation³⁰.

The structure of ZP3 ZP-N offers a first glimpse into the basis of mammalian fertilization at atomic resolution. Because many pathological mutations in other ZP domain proteins fall within ZP-N, this work will not only be relevant for reproductive medicine, but also for severe human diseases such as non-syndromic deafness or renal and vascular disorders.

METHODS SUMMARY

Proteins were co-expressed with a leaderless version of disulphide isomerase DsbC and TrxA mutant Gly33Pro/Pro34Tyr in *E. coli* Origami B (DE3)⁶. Monomeric fusions purified by affinity and size-exclusion chromatography were crystallized by vapour diffusion with sodium acetate (crystal form I), PEG 6000 (crystal form II) and PEG 6000/ZnCl₂ (crystal form III) as precipitants (Supplementary Fig. 1g–i). Diffraction data was collected at beamlines BM14 (crystal forms I and II) and ID23-1 (crystal form III) of the European Synchrotron Radiation Facility (ESRF), Grenoble. Crystal form I was refined to 2.9 Å resolution in space group *I*222, with a crystallographic *R*-factor (*R*_{work}) of 19.0% and a free *R*-factor (*R*_{free}) of 22.7% (Supplementary Table 1). Triclinic crystal form II was refined to 2.3 Å resolution with *R*_{work} = 20.1% and *R*_{free} = 22.6% (Supplementary Table 1 and Supplementary Fig. 2). Crystal form III was refined to 3.1 Å resolution in space group *P*₂₁₂₁₂, with *R*_{work} = 20.8% and *R*_{free} = 25.6% (Supplementary Table 1). The final models for crystal forms I and II, that have essentially identical structure (average ZP-N main chain root mean square (r.m.s.) distance = 0.5 Å) as well as very similar cell dimensions and crystal packing (Supplementary Fig. 3), include all ZP-N residues and have excellent stereochemistry, with no Ramachandran plot outliers. Crystal form III has also similar cell dimensions and structure (r.m.s. distance = 0.8 Å), but lacks electron density for the FG loop and differs in packing (Fig. 3c).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 September; accepted 29 October 2008.

1. Nixon, B., Aitken, R. J. & McLaughlin, E. A. New insights into the molecular mechanisms of sperm-egg interaction. *Cell. Mol. Life Sci.* **64**, 1805–1823 (2007).
2. Wassarman, P. M. Zona pellucida glycoproteins. *J. Biol. Chem.* **283**, 24285–24289 (2008).

3. Bork, P. & Sander, C. A large domain common to sperm receptors (Zp2 and Zp3) and TGF- β type III receptor. *FEBS Lett.* **300**, 237–240 (1992).
4. Jovine, L., Qi, H., Williams, Z., Litscher, E. & Wassarman, P. M. The ZP domain is a conserved module for polymerization of extracellular proteins. *Nature Cell Biol.* **4**, 457–461 (2002).
5. Jovine, L., Darie, C. C., Litscher, E. S. & Wassarman, P. M. Zona pellucida domain proteins. *Annu. Rev. Biochem.* **74**, 83–114 (2005).
6. Jovine, L., Janssen, W. G., Litscher, E. S. & Wassarman, P. M. The PLAC1-homology region of the ZP domain is sufficient for protein polymerisation. *BMC Biochem.* **7**, 11 (2006).
7. Llorca, O., Trujillo, A., Blanco, F. J. & Bernabeu, C. Structural model of human endoglin, a transmembrane receptor responsible for hereditary hemorrhagic telangiectasia. *J. Mol. Biol.* **365**, 694–705 (2007).
8. Boja, E. S., Hoodbhoy, T., Fales, H. M. & Dean, J. Structural characterization of native mouse zona pellucida proteins using mass spectrometry. *J. Biol. Chem.* **278**, 34189–34202 (2003).
9. Kanai, S. *et al.* Disulfide linkage patterns of pig zona pellucida glycoproteins ZP3 and ZP4. *Mol. Reprod. Dev.* **75**, 847–856 (2008).
10. Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* **12**, 1313–1322 (2003).
11. Halaby, D. M., Poupon, A. & Mornon, J. The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng.* **12**, 563–571 (1999).
12. Yang, Y. *et al.* Structural basis for dimerization of ICAM-1 on the cell surface. *Mol. Cell* **14**, 269–276 (2004).
13. Doren, S. *et al.* Incorporation of mouse zona pellucida proteins into the envelope of *Xenopus laevis* oocytes. *Dev. Genes Evol.* **209**, 330–339 (1999).
14. ten Dijke, P., Goumans, M. J. & Pardali, E. Endoglin in angiogenesis and vascular diseases. *Angiogenesis* **11**, 79–89 (2008).
15. Legan, P. K. *et al.* A deafness mutation isolates a second role for the tectorial membrane in hearing. *Nature Neurosci.* **8**, 1035–1042 (2005).
16. Callebaut, I., Mornon, J. P. & Monget, P. Isolated ZP-N domains constitute the N-terminal extensions of Zona Pellucida proteins. *Bioinformatics* **23**, 1871–1874 (2007).
17. Lindsay, L. L. & Hedrick, J. L. Proteolysis of *Xenopus laevis* egg envelope ZPA triggers envelope hardening. *Biochem. Biophys. Res. Commun.* **324**, 648–654 (2004).
18. Tian, J., Gong, H. & Lennarz, W. J. *Xenopus laevis* sperm receptor gp69/64 glycoprotein is a homolog of the mammalian sperm receptor ZP2. *Proc. Natl Acad. Sci. USA* **96**, 829–834 (1999).
19. Wassarman, P. M. & Mortillo, S. Structure of the mouse egg extracellular coat, the zona pellucida. *Int. Rev. Cytol.* **130**, 85–110 (1991).
20. Tsubamoto, H. *et al.* Expression of recombinant human zona pellucida protein 2 and its binding capacity to spermatozoa. *Biol. Reprod.* **61**, 1649–1654 (1999).
21. Hasegawa, A., Hamada, Y., Shigeta, M. & Koyama, K. Contraceptive potential of synthetic peptides of zona pellucida protein (ZPA). *J. Reprod. Immunol.* **53**, 91–98 (2002).
22. Sun, W., Lou, Y. H., Dean, J. & Tung, K. S. A contraceptive peptide vaccine targeting sulfated glycoprotein ZP2 of the mouse zona pellucida. *Biol. Reprod.* **60**, 900–907 (1999).
23. Bleil, J. D., Beall, C. F. & Wassarman, P. M. Mammalian sperm-egg interaction: fertilization of mouse eggs triggers modification of the major zona pellucida glycoprotein, ZP2. *Dev. Biol.* **86**, 189–197 (1981).
24. Rankin, T. L. *et al.* Fertility and taxon-specific sperm binding persist after replacement of mouse sperm receptors with human homologs. *Dev. Cell* **5**, 33–43 (2003).
25. Aricescu, A. R. & Jones, E. Y. Immunoglobulin superfamily cell adhesion molecules: zippers and signals. *Curr. Opin. Cell Biol.* **19**, 543–550 (2007).
26. Familiari, G., Heyn, R., Relucenti, M. & Sathananthan, H. Structural changes of the zona pellucida during fertilization and embryo development. *Front. Biosci.* **13**, 6730–6751 (2008).
27. Makalowski, W. & Boguski, M. S. Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc. Natl Acad. Sci. USA* **95**, 9407–9412 (1998).
28. Swanson, W. J., Yang, Z., Wolfner, M. F. & Aquadro, C. F. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc. Natl Acad. Sci. USA* **98**, 2509–2514 (2001).
29. Turner, L. M. & Hoekstra, H. E. Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice (*Peromyscus*). *Mol. Biol. Evol.* **23**, 1656–1669 (2006).
30. Galindo, B. E., Vacquier, V. D. & Swanson, W. J. Positive selection in the egg receptor for abalone sperm lysin. *Proc. Natl Acad. Sci. USA* **100**, 4639–4643 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgments This work was supported by Karolinska Institutet, the Swedish Research Council (grant 2005-5102) and European Community (Marie Curie ERG 31055). We thank P. Nordlund and G. Schneider for access to the Stockholm SGC robotic crystallization facility and beamtime; H. Belrhali, J. McCarthy, R. Ravelli and M. Walsh for assistance at ESRF; P. Afonine, R. Grosse-Kunstleve and P. Zwart for help with phenix.refine; G. Murshudov for help with REFMAC; C. Chothia, F. Cotelli, J.-Å. Gustafsson, R. Herbst-Irmer, A. Kohl, R. Ladenstein, M. Letarte, E. Litscher, E. Morgunova, A. Murzin, K. Nagai, L. Nilsson, D. Rhodes, R. Toftgård and P. Wassarman for discussions and comments.

Author Contributions M.M., L.H. and L.J. generated constructs and performed protein expression and purification. T.S. performed mass spectrometric analysis. S.B. carried out molecular dynamics simulations. L.J. and M.M. crystallized proteins, determined structures and wrote the manuscript.

Author Information Atomic coordinates and structure factors have been deposited with the Protein Data Bank under accession codes 3D4C (crystal form I), 3D4G (crystal form II) and 3EF7 (crystal form III). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.J. (luca.jovine@ki.se).

METHODS

DNA constructs. A PCR fragment encoding ZP3 ZP-N was cloned into vectors pLMBP4c (ref. 6) or pLMBP6c to generate constructs expressing cytoplasmic MBP Glu385Ala/Lys388Ala/Asp389Ala/Arg393Asn (ref. 31), fused via an Ala-Ala linker to C- or N-terminally histidine-tagged ZP3 ZP-N. Point mutations were introduced using the QuickChange protocol (Stratagene). All constructs were verified by DNA sequencing.

Protein expression and purification. Proteins were expressed essentially as described⁶. Soluble fusions were captured by immobilized-metal affinity chromatography using HisTrap HP columns (GE Healthcare), and monomeric protein was eluted separately from oligomeric and aggregated species with a 2 column volume linear gradient to 86 mM imidazole ($\sim 9.4 \text{ mS cm}^{-1}$), followed by an 18 column volume gradient to 500 mM imidazole (Supplementary Fig. 1b, d, e). Concentrated peak fractions from the first gradient were incubated with 10 mM D-(+)-maltose, further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 prep grade column (GE Healthcare) (Supplementary Fig. 1c–e) and concentrated to 15 mg ml^{-1} . The identity of the purified product and the presence of native disulphide bond connectivity were confirmed by mass spectrometry. Immunoblotting experiments were performed as previously described⁶.

Mass spectrometry. Samples were prepared as described⁶ and analysed by MALDI-TOF and MALDI-TOF/TOF (Ultraflex II, Bruker Daltonics). Spectra were annotated with the Flexanalysis software (Bruker Daltonics).

Crystallization and diffraction data collection. Crystal form I of MBP-AAA-ZP-N-6His grew in 3–10 days at 20°C in 1.0 M sodium acetate, 100 mM Na-HEPES, pH 7.5, 50 mM CdSO_4 . Macroseeded specimens ($\sim 250 \times 200 \times 100 \mu\text{m}$; Supplementary Fig. 1g) were cryoprotected by gradual transfer into 1.5 M magnesium acetate, 100 mM Na-HEPES, pH 7.5, 10 mM Tris-HCl, pH 7.2, 50 mM CdSO_4 , 1 mM D-(+)-maltose, mounted in cryoloops and flash frozen in liquid nitrogen. They diffracted to 2.9 \AA at synchrotron and belonged to space group *I*222, with one fusion protein per asymmetric unit (Supplementary Table 1). Essentially identical crystals could also be grown in the same condition using an N-terminally tagged fusion construct (6His-MBP-AAA-ZP-N); however, because the diffraction quality of these crystals was not superior to those of MBP-AAA-ZP-N-6His, they were not pursued further.

Crystal form II of MBP-AAA-ZP-N-6His grew over the course of 3 weeks in 8–24% PEG 6000, 100 mM Na-HEPES, pH 7.0–8.0, 125–250 mM CaCl_2 at 20°C . Crystals ($\sim 250 \times 125 \times 125 \mu\text{m}$; Supplementary Fig. 1h) were cryoprotected by stepwise transfer into a solution equivalent to the mother liquor plus 20% PEG 400. They diffracted to a Bragg spacing of 2.3 \AA at synchrotron, with eight molecules per asymmetric unit in a primitive triclinic lattice (Supplementary Table 1). Also in the case of this crystal form, crystals of construct 6His-MBP-AAA-ZP-N could be grown in related conditions that had the same symmetry and very similar unit cell parameters. A 2.7 \AA resolution data set ($R_{\text{sym}} = 3.7$ (43.2)) was collected to assess whether the position of the histidine tag within the fusion constructs had any effect on the structure of ZP-N.

Crystal form III of MBP-AAA-ZP-N-6His grew into feather-shaped aggregates in 2 days in 12% PEG 6000, 100 mM Tris-HCl, pH 8.2, 150 mM CaCl_2 , 2.5 mM ZnCl_2 at 20°C (Supplementary Fig. 1i). Pieces of a crushed aggregate were cryoprotected in 12% PEG 6000, 100 mM Na-HEPES, pH 7.8, 150 mM CaCl_2 , 5 mM ZnCl_2 supplemented with 20% PEG 400 and screened at synchrotron for presence of a single lattice. A 3.1 \AA resolution data set was obtained, which suggested that the protein crystallized in space group *P*₂₁₂, with two molecules per asymmetric unit. All data sets were processed with the XDS package³².

Structure determination. Crystal form I: The structure of MBP-AAA-ZP-N-6His was first solved in space group *I*222 by molecular replacement with Phaser³³, using an ensemble of selected MBP structures as search model. Correctness of the solution was confirmed by strong difference electron density ($>6\sigma$) for a molecule of D-(+)-maltose (coordinates for which were not included in the search ensemble) within the MBP ligand-binding pocket. After statistical phase improvement in Pirate³⁴, clear density could also be observed for fragments of ZP-N which, as expected, were positioned on the C-terminal side of the MBP moiety. The ZP-N model was built manually with Coot³⁵ and the backrub tool of KING³⁶, using 3.2 \AA resolution SigmaA-weighted difference Fourier maps³⁷, composite annealed omit maps calculated with CNS³⁸ using a starting temperature of 4,500 K to minimize model bias, and X-ray holographic image reconstruction in Eden³⁹. The position of the two conserved disulphide bridges of ZP-N was confirmed by sulphur anomalous signal in phased difference maps generated using 5.9-fold redundant low resolution data (Fig. 2). The structure was refined with phenix.refine⁴⁰ using a maximum likelihood (ML) target. Bulk solvent correction was applied throughout, and Translation/Libration/Screw refinement of B factors was performed during final rebuilding cycles, on the basis of TLSMD⁴¹ analysis of individually refined atomic displacement parameters.

Crystal form II: A set of coordinates partially refined against the *I*222 data was used to solve the high resolution structure of MBP-AAA-ZP-N-6His by molecular replacement with Phaser. Initial refinement was complicated by the presence of strong pseudo-orthorhombic symmetry, including a translational non-crystallographic symmetry (NCS) component. However, a series of parallel test refinement jobs in all possible space groups clearly indicated that this crystal form belongs to the triclinic system and allowed structure determination to proceed. The structure was refined in space group *P*₁ with phenix.refine and REFMAC⁴², essentially as described above; partial NCS restraints were also applied, except in the final cycles of refinement. NCS editing features of Coot were used during rebuilding, but molecules within the asymmetric unit were also individually rebuilt at two different stages of refinement. FG loops were built iteratively on the basis of traces generated with Xpleo⁴³; solvent structure was assembled by phenix.refine and manually edited in Coot. Comparison of the final model (Supplementary Fig. 2) to partially refined coordinates of 6His-MBP-AAA-ZP-N ($R_{\text{work}} = 19.3\%$, $R_{\text{free}} = 26.0\%$) revealed essentially no difference between ZP-N moieties (average main chain r.m.s. distance excluding FG loop residues = 0.13 \AA), indicating that the structure of the domain is not influenced by presence of a C-terminal histidine tag.

Crystal form III: The structure was solved by molecular replacement with Phaser, using chain A of the refined *P*₁ coordinates as search model. Ordered Zn^{2+} ions were identified in phased anomalous electron density maps and used to obtain experimental phase information using the EP_SAD mode of Phaser (figure of merit = 0.280 (0.132)). Refinement against a 7.6-fold redundant data set collected at the Zn K-edge ($\lambda = 1.28235 \text{ \AA}$) was carried out with phenix.refine, using a phased maximum likelihood (MLHL) target and anomalous scattering factors ($f' = -8.21$, $f'' = 4.75$) derived from X-ray fluorescence spectra with CHOOCH⁴⁴. Because visual inspection of initial maps revealed that one of the two molecules within the asymmetric unit (chain B) was relatively disordered (presumably due to the lack of β -strand E'-mediated contacts), no NCS restraints were used in this case.

Models were validated with MolProbity⁴⁵ and ValigURL⁴⁶. The side chains of ZP3 residues Glu 45 (*P*₂₁₂ only); Glu 48 (*P*₁ chain C, *P*₂₁₂ chain A); Lys 43; Leu 59 (*P*₂₁₂ chain B); Lys 64; Gln 67 (*P*₂₁₂); Leu 73 (*P*₂₁₂ chain B); Ser 75 (*I*222); Glu 76; Gln 79 (*P*₂₁₂ chain B); Arg 81 and Arg 90 (*P*₂₁₂); Arg 101 (*P*₂₁₂ chain B); Lys 106; Asp 107, Arg 132 and Arg 140 (*P*₂₁₂ chain B) are partially disordered. Side chains of FG loop residues Arg 120–Arg 129 are poorly resolved, and backbone density for Ser 123–Ser 126 (*I*222) or Pro 121–Ser 126 (*P*₁) is weak; nevertheless, the latter regions could be modelled because their position is constrained by the better defined residues that flank them within the loop. No electron density could be observed for FG loop residues Pro 119–Asn 131 in the *P*₂₁₂ map.

Sequence and structure analysis. Structure-based multiple sequence alignments were assembled using STRAP⁴⁷ and 3D-Coffee⁴⁸, starting from iteratively optimized pairwise alignments generated by MODELLER⁴⁹. Residues in Fig. 1c are colour-coded as follows: brown, hydrophobic; purple, aromatic; orange, [Pro, Gly]; white, Cys; green, polar; blue, positively charged; red, negatively charged. The 70% threshold consensus sequences for ZP3 homologues with $\leq 90\%$ identity (ZP3 cons.) and for Pfam ZP domain family seed sequences (Pfam cons.) were calculated with Consensus⁵⁰; lowercase characters correspond to the following amino acid sets: a/aromatic = [Phe, Tyr, Trp, His]; l/aliphatic = [Ile, Val, Leu]; h/hydrophobic = [Phe, Tyr, Trp, His, Ile, Val, Leu, Ala, Gly, Met, Cys, Lys, Arg, Thr]; +/positive = [His, Lys, Arg]; –/negative = [Asp, Glu]; c/charged = [His, Lys, Arg, Asp, Glu]; p/polar = [His, Lys, Arg, Asp, Glu, Gln, Asn, Ser, Thr, Cys]; o/alcohol = [Ser, Thr]; u/tiny = [Gly, Ala, Ser]; s/small = [Gly, Ala, Ser, Val, Thr, Asp, Asn, Pro, Cys]; t/turnlike = [Gly, Ala, Ser, His, Lys, Arg, Asp, Glu, Gln, Asn, Ser, Thr, Cys]; (.) = any residue; uppercase letters indicate the specific amino acids with the same one-letter code. Evolutionary conservation was analysed with ConSurf⁵¹, using alignments of non-redundant sequence data sets generated with 90% identity threshold. Interfaces were studied with PISA⁵², cavities were analysed with CAVER⁵³ and Poisson-Boltzmann electrostatic calculations were performed using APBS⁵⁴. Structural similarity was evaluated with SSM⁵⁵ and Dali⁵⁶. Protein–protein interaction sites were predicted with PINUP⁵⁷, ISIS⁵⁸, ProMate⁵⁹, SPIDER⁶⁰ and InterProSurf⁶¹. Figures were prepared using PyMOL⁶².

Homology recognition and modelling of isolated ZP-N domains. Non-ZP domain ZP-Ns were identified using Fugue⁶³ and a HOMSTRAD database⁶⁴ that consisted of 8,724 folds and included a profile for ZP-N. MODELLER-generated homology models of the hits were then used to generate extra profiles for further scanning of ZP domain protein sequences, and the procedure was repeated until no more copies of ZP-N could be found. Using this method, all ZP-N repeats of mouse ZP2 were identified with $>95\%$ confidence. Parallel searches were also performed with a profile hidden Markov model created using HMMER⁶⁵, on the basis of ZP3 ZP-N secondary structure and surface accessibility information.

Modelling of the N-terminal fragment of ZP2. Individual models of ZP-N domains 1–4 of mouse ZP2 (ZP-N1, Gly 47–Gln 138; ZP-N2, Glu 146–Asp 258; ZPN-3, Leu 259–Ser 357; ZP-N4, Ile 361–Arg 462) were created with MODELLER and merged into a single chain using Coot. Relative orientation of the domains was constrained by the short loop sequences connecting adjacent ZP-Ns. This was particularly evident in the case of ZP-N2 and ZP-N3, in which, owing to the absence of an interdomain linker, the first β -strand of ZP-N3 is positioned so that it extends in the same direction as the last β -strand of ZP-N2; as a result, the BC and FG loops of ZP-N3 wrap around the lower part of ZP-N2. A similar arrangement is observed in the case of ZP-N4 and ZP-N3, respectively. Relative positioning of ZP-N1 and ZP-N2 depends on interactions between their longer intradomain linker and both the BC and FG loops of ZP-N2; the tip of the former loop, which is particularly long owing to insertion of the post-fertilization cleavage site, also directly contacts the bottom part of ZP-N1. Coordinates were energy minimized using CHARMM⁶⁶, resulting in a final model with 96% of the residues in allowed Ramachandran plot regions. Accessibility of sites experimentally determined to be N-glycosylated⁸ was assessed with GlyProt⁶⁷.

Molecular dynamics simulations. Coordinates of ZP3 residues from chain A of the P1 structure were used to set up a wild-type model of ZP-N in CHARMM. A model of mutant Tyr111Cys was constructed from the wild-type coordinates by modifying the Tyr 111 side chain with CHARMM and set up with two alternative disulphide bond connectivities, one in which the mutated residue was free (Tyr111Cys*) and the other one where Cys111 was disulphide-bonded to Cys₁₃₉, leaving Cys₁₄₆ unpaired (Tyr111Cys^X). Models were solvated using TIP3P water⁶⁸ in a 65×40×40 Å box. The system was minimized with the steepest descent and ABNR method, initially with C α atom constraints (force constant 20/50 kcal mol⁻¹ Å⁻²) and then without constraints. A standard simulation protocol was used⁶⁹ with temperature set to 300 K and allowed to vary within ± 10 K, as well as periodic boundary conditions and weak constraints on C α atoms (1 kcal mol⁻¹ Å⁻²) to maintain biochemical properties. Molecular dynamics simulations were performed for 10 ns and systems were equilibrated for 400 ps before data collection started.

Mammalian cell culture and transient expression experiments. Complementary DNA fragments encoding full-length wild-type mouse ZP3 and mutants Tyr111Cys and Tyr111Phe were cloned into vector pSI (Promega) and transfected into CHO-K1 cells (American Type Culture Collection) as previously described⁴. Immunoblotting was performed with anti-ZP3 polyclonal antibody 8818 (1:10,000 dilution; Pocono Rabbit Farm) and peroxidase-conjugated AffiniPure goat anti-rabbit IgG (H+L) (1:10,000; Jackson ImmunoResearch Europe).

31. Center, R. J. *et al.* Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci.* **7**, 1612–1619 (1998).
32. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
33. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
34. Cowtan, K. General quadratic functions in real and reciprocal space and their application to likelihood phasing. *Acta Crystallogr. D* **56**, 1612–1621 (2000).
35. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Davis, I. W., Arendall, W. B. III, Richardson, D. C. & Richardson, J. S. The backbone motion: how protein backbone shrugs when a sidechain dances. *Structure* **14**, 265–274 (2006).
37. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).
38. Brünger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
39. Somoza, J. R. *et al.* Holographic methods in X-ray crystallography. IV. A fast algorithm and its application to macromolecular crystallography. *Acta Crystallogr. A* **51**, 691–708 (1995).
40. Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. The Phenix refinement framework. *CCP4 Newsletter* **42**, contribution 8 (2005).
41. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* **62**, 439–450 (2006).
42. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
43. van den Bedem, H., Lotan, I., Latombe, J. C. & Deacon, A. M. Real-space protein-model completion: an inverse-kinematics approach. *Acta Crystallogr. D* **61**, 2–13 (2005).
44. Evans, G. & Pettifer, R. F. CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst.* **34**, 82–86 (2001).
45. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
46. Kleywegt, G. J. & Harris, M. R. ValLigURL: a server for ligand-structure comparison and validation. *Acta Crystallogr. D* **63**, 935–938 (2007).
47. Gille, C. & Frommel, C. STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics* **17**, 377–378 (2001).
48. O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D. G. & Notredame, C. 3DCoffee: combining protein sequences and structures within multiple sequence alignments. *J. Mol. Biol.* **340**, 385–395 (2004).
49. Fiser, A. & Sali, A. Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol.* **374**, 461–491 (2003).
50. Bork, P. Consensus. <<http://coot.embl.de/Alignment/consensus.html>>.
51. Landau, M. *et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 (2005).
52. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
53. Petrek, M. *et al.* CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **7**, 316 (2006).
54. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).
55. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D* **60**, 2256–2268 (2004).
56. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
57. Liang, S., Zhang, C., Liu, S. & Zhou, Y. Protein binding site prediction using an empirical scoring function. *Nucleic Acids Res.* **34**, 3698–3707 (2006).
58. Ofra, Y. & Rost, B. Protein-protein interaction hotspots carved into sequences. *PLOS Comput. Biol.* **3**, e119 (2007).
59. Neuvirth, H., Raz, R. & Schreiber, G. ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J. Mol. Biol.* **338**, 181–199 (2004).
60. Porollo, A. & Meller, J. Prediction-based fingerprints of protein-protein interactions. *Proteins* **66**, 630–645 (2007).
61. Negi, S. S. & Braun, W. Statistical analysis of physical-chemical properties and prediction of protein-protein interfaces. *J. Mol. Model.* **13**, 1157–1167 (2007).
62. DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific, 2002).
63. Shi, J., Blundell, T. L. & Mizuguchi, K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310**, 243–257 (2001).
64. Stebbings, L. A. & Mizuguchi, K. HOMSTRAD: recent developments of the Homologous Protein Structure Alignment Database. *Nucleic Acids Res.* **32**, D203–D207 (2004).
65. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).
66. Brooks, B. R. *et al.* CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
67. Böhne-Lang, A. & von der Lieth, C. W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.* **33**, W214–W219 (2005).
68. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
69. Burendahl, S., Treuter, E. & Nilsson, L. Molecular dynamics simulations of human LRH-1: the impact of ligand binding in a constitutively active nuclear receptor. *Biochemistry* **47**, 5205–5215 (2008).

LETTERS

The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor

Sarah E. Ewald¹, Bettina L. Lee¹, Laura Lau¹, Katherine E. Wickliffe¹, Guo-Ping Shi², Harold A. Chapman³ & Gregory M. Barton¹

Mammalian Toll-like receptors (TLRs) 3, 7, 8 and 9 initiate immune responses to infection by recognizing microbial nucleic acids^{1,2}; however, these responses come at the cost of potential autoimmunity owing to inappropriate recognition of self nucleic acids³. The localization of TLR9 and TLR7 to intracellular compartments seems to have a role in facilitating responses to viral nucleic acids while maintaining tolerance to self nucleic acids, yet the cell biology regulating the transport and localization of these receptors remains poorly understood^{4–6}. Here we define the route by which TLR9 and TLR7 exit the endoplasmic reticulum and travel to endolysosomes in mouse macrophages and dendritic cells. The ectodomains of TLR9 and TLR7 are cleaved in the endolysosome, such that no full-length protein is detectable in the compartment where ligand is recognized. Notably, although both the full-length and cleaved forms of TLR9 are capable of binding ligand, only the processed form recruits MyD88 on activation, indicating that this truncated receptor, rather than the full-length form, is functional. Furthermore, conditions that prevent receptor proteolysis, including forced TLR9 surface localization, render the receptor non-functional. We propose that ectodomain cleavage represents a strategy to restrict receptor activation to endolysosomal compartments and prevent TLRs from responding to self nucleic acids.

The subcellular localization of TLR9 and TLR7 is highly regulated. Although ligand recognition occurs in endolysosomes, it has been reported that most, if not all, TLR9 resides in the endoplasmic reticulum (ER) of resting cells^{5,6}. Curiously, analyses of N-linked glycosylation on TLR9 indicates that the protein does not travel through the Golgi, even in activated cells^{5,6}. These observations have led to a model in which TLR9 is recruited directly to endolysosomes from the ER, bypassing the classical secretory pathway⁵. However, general mechanistic details accounting for such unusual cell biology are lacking, and the specific details of how TLR9 transport is regulated between these compartments remain largely unknown.

To separate ER-resident TLR9 from receptors in the endolysosome, we purified phagosomes from cells fed latex beads. To track TLR9, we used macrophage-like RAW264 cells stably expressing TLR9 with a carboxy-terminal haemagglutinin (HA) tag (RAW-TLR9 cells). We have previously shown that this epitope tag does not affect TLR9 function⁴. As expected, phagosome preparations were enriched for the lysosomal protein LAMP1 whereas they were devoid of markers of other intracellular organelles and the cytosol (Fig. 1a). Surprisingly, we did not observe any TLR9 at the expected molecular mass (150 kDa). Instead, an approximately 80 kDa protein was highly enriched in the phagosomal preparation (Fig. 1a). The presence of ligand did not alter receptor translocation, as we could

not detect full-length TLR9 in phagosomes containing latex beads coated with CpG oligonucleotides (Fig. 1b), despite the fact that these beads induced TLR9 signalling (Supplementary Fig. 1a). In contrast, the 80 kDa protein was evident in phagosomes at the earliest time point examined (10 min) and remained equally abundant over 4 h (Fig. 1b).

To demonstrate directly that the 80 kDa band in our phagosome preparations represented a cleaved form of TLR9, we performed pulse–chase analysis. The 80 kDa band was first detectable after 3 h of chase, concomitant with the onset of full-length receptor degradation (Fig. 1c). The 80 kDa band continued to accumulate in parallel with the degradation of full-length TLR9, peaking at 6 h when the full-length protein was no longer detectable. These data formally demonstrate that the 80 kDa protein is a cleavage product of full-length TLR9 consisting of approximately half of the ectodomain, the transmembrane domain and the entire cytoplasmic domain. On the basis of molecular mass, we estimate that cleavage occurs within or distal to leucine-rich repeat (LRR) 14. Structural modelling of TLR9 suggests that the cleaved receptor adopts an abbreviated ‘horseshoe’ fold encompassing a region that corresponds to residues implicated in ligand binding by TLR3 (refs 7, 8) (Fig. 1d). A C-terminal cleavage product of TLR9 was also present in macrophage colony-stimulating factor (M-CSF)-derived macrophages, granulocyte–macrophage (GM)-CSF-derived dendritic cells, and Flt3L-derived plasmacytoid dendritic cells, indicating that similar processing events occur in many immunologically relevant cell types (Fig. 1e). We observed a similar cleavage product of TLR7, but not of TLR4, indicating that receptor processing may be a general feature shared by nucleic-acid-sensing TLRs (Fig. 1f).

We next sought to determine where within the cell TLR9 cleavage occurs. Although N-linked sugars on full-length TLR9 were sensitive to endoglycosidase H (EndoH), most of the sugars on the cleaved form of TLR9 were resistant to EndoH, indicating that the truncated form of the receptor has passed through the Golgi (Fig. 2a). A similar glycosylation pattern was observed for TLR7 (Supplementary Fig. 1b). In addition, a small fraction of full-length TLR9 consistently migrated slightly slower than most of the full-length protein (Fig. 2b, see asterisk). Notably, this high-molecular-mass band was almost entirely EndoH resistant, in contrast to the rest of the full-length TLR9 protein, indicating that this small pool of full-length protein has also passed through the Golgi. We also observed this high-molecular-mass, EndoH-resistant band in pulse–chase experiments (asterisks in Fig. 1c and Supplementary Fig. 1c). The band appeared after 1.5 h of chase, immediately preceding formation of the cleavage product. Furthermore, the intensity of the high-molecular-mass band appeared to match the intensity of the accumulated cleavage product.

¹Division of Immunology & Pathogenesis, Department of Molecular and Cell Biology, University of California, Berkeley, 405 Life Sciences Addition, Berkeley, California 94720-3200, USA. ²Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, NRB-7, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. ³Department of Medicine, and The Cardiovascular Research Institute, University of California, San Francisco, Box 0111, San Francisco, California 94143, USA.

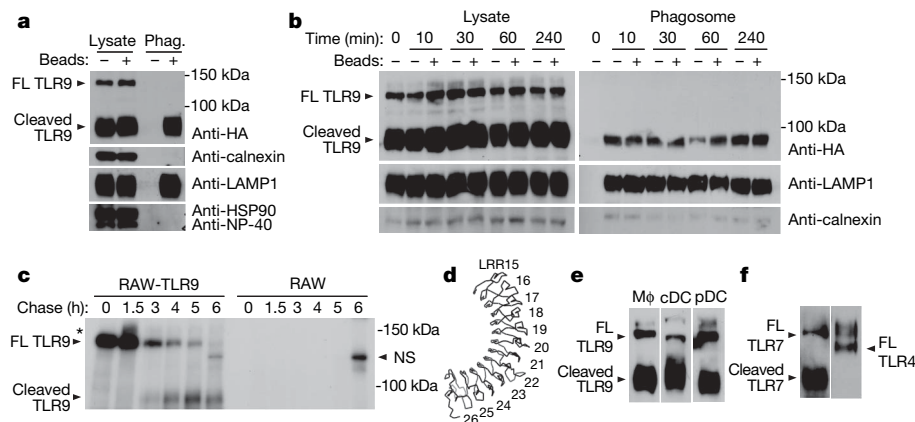


Figure 1 | A truncated version of TLR9 is exclusively present in phagosomes. **a**, Lysates or phagosome preparations from RAW-TLR9 cells fed latex beads were separated by SDS-PAGE and probed with antibodies specific for the indicated proteins. Full-length TLR9 (FL TLR9) and the truncated protein (cleaved TLR9) are indicated. **b**, Analysis of phagosomes from RAW-TLR9 cells fed unconjugated latex beads or beads conjugated to stimulatory CpG oligonucleotides. Lysates or purified phagosomes from the cells were separated by SDS-PAGE and probed with antibodies against the indicated proteins. **c**, The truncated form of TLR9 is a product of the full-length protein. Pulse-chase analysis of RAW-TLR9 cells or control cells (RAW). An asterisk indicates the high-molecular-mass full-length TLR9 band. NS indicates a nonspecific band present in both RAW-TLR9 and

control RAW cells. **d**, Modelled structure of the truncated ectodomain (residues 477–790) of TLR9. Leucine-rich repeats (LRR) 15–26 are labelled. **e**, The TLR9 cleavage product is present in multiple cell types. Haematopoietic stem cells were transduced with a retrovirus encoding TLR9–HA and then differentiated into macrophages with M-CSF (Mφ), dendritic cells with GM-CSF (cDC), or plasmacytoid dendritic cells with Flt3L (pDC). **f**, Lysates from RAW cells expressing TLR4–HA or TLR7–HA were probed with anti-HA antibody. Note that TLR4 is not cleaved. The upper band of TLR4 is due to glycosylation. Unless indicated otherwise, the data presented in each panel are representative of at least three independent experiments.

Collectively, these data suggest that this high-molecular-mass band represents the small fraction of TLR9 that exits the ER, passes through the Golgi, and is cleaved. In support of this model, cells treated with brefeldin A (BFA), a reagent that disrupts the Golgi, contained neither the high-molecular-mass band (1.5-h chase) nor the 80-kDa cleavage product (5-h chase) (Fig. 2c). Analysis of TLR9 localization by immunofluorescence microscopy corroborated our biochemical findings (Supplementary Figs 2 and 3).

To address the mechanism responsible for TLR9 sorting from the ER, we examined UNC93B1 (refs 9–11). Mice with a deficiency in UNC93B1 no longer respond to TLR3, TLR7 and TLR9 ligands¹¹, and UNC93B1 can associate with TLR3, TLR7 and TLR9 which may facilitate transport to endolysosomes^{9,10}. We generated RAW cells in which *Unc93b1* mRNA was knocked down by retrovirally encoded short hairpin RNA (shRNA) (Fig. 3a). As expected, these cells exhibited reduced TNF- α production in response to CpG and R848 relative to control cells (Fig. 3b). Notably, the TLR9 cleavage product was greatly reduced in UNC93B1 knockdown cells, although full-length TLR9 was expressed at levels comparable to control cells (Fig. 3c). The high molecular mass, full-length TLR9 band was also absent in UNC93B1 knockdown cells (Fig. 3c), and TLR9 was not present in phagosomes purified from UNC93B1 knockdown cells (Fig. 3d). These data demonstrate that UNC93B1 is required for TLR9 exit from the ER. To test whether UNC93B1 is also sufficient for this exit, we overexpressed UNC93B1 in mouse embryonic fibroblasts (MEFs), which signal poorly in response to CpG oligonucleotides. Whereas we detected very little TLR9 cleavage in control cells, simply overexpressing UNC93B1 in MEFs resulted in a 75-fold increase in the 80-kDa band (Fig. 3e). The increase in cleaved TLR9 corresponded with a robust increase in responsiveness to CpG oligonucleotides (Fig. 3f). Thus, UNC93B1 seems to control the functional pool of TLR9 by regulating exit from the ER. Furthermore, the fraction of TLR9 that exits the ER and is cleaved seems to be responsible for TLR9 signalling.

We next attempted to block TLR9 cleavage by inhibiting various cellular proteases. Bafilomycin A1, an inhibitor of the vacuolar ATPase, blocks endolysosomal acidification as well as signalling by TLR9, TLR7 and TLR3 (ref. 12). Pulse-chase analysis of RAW-TLR9 cells treated with bafilomycin A1 revealed that TLR9 cleavage was

completely blocked when endolysosomal acidification was inhibited (Fig. 4a). Blocking TLR9 proteolysis with bafilomycin A1 also resulted

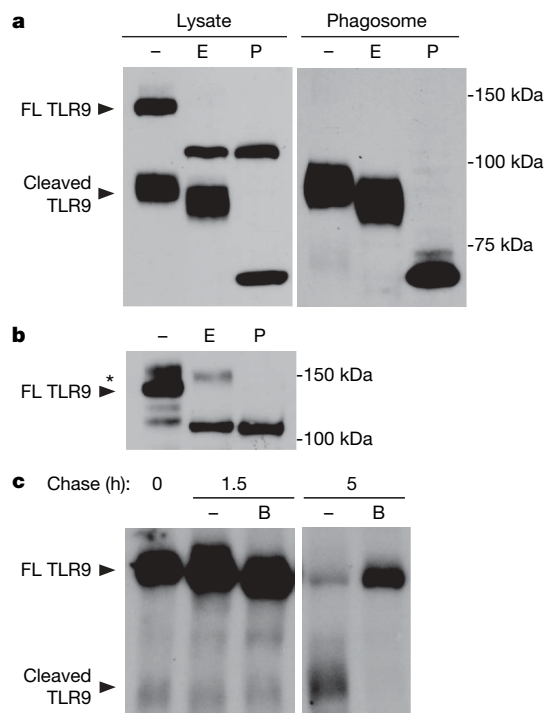


Figure 2 | TLR9 exits the ER and travels through the Golgi before cleavage. **a**, Cleaved TLR9 is EndoH-resistant. Proteins from lysates or purified phagosomes from RAW-TLR9 cells were immunoprecipitated with anti-HA resin and treated with EndoH (E), PNGaseF (P), or left untreated (–). **b**, A longer exposure of TLR9 immunoprecipitated from RAW-TLR9 cells and subjected to analysis as described in **a**. An asterisk indicates the high-molecular-mass full-length TLR9 band that is EndoH-resistant. **c**, TLR9 cleavage requires passage through the Golgi. Pulse-chase analysis was performed, at the indicated chase time points, on RAW-TLR9 cells treated with brefeldin A (B) or vehicle (–). The data presented in each panel are representative of at least three independent experiments.

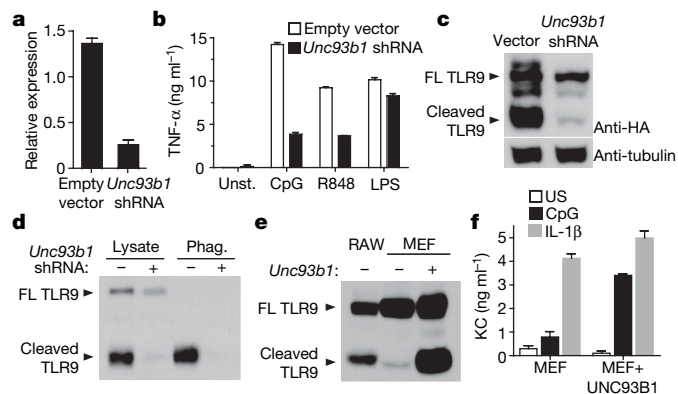


Figure 3 | *UNC93B1* is required for TLR9 ER exit and cleavage.

a, Expression of *Unc93b1* transcript was measured by quantitative RT-PCR in RAW cells transduced with a control retrovirus or a retrovirus encoding an shRNA specific for *Unc93b1*. **b**, *Unc93b1* shRNA RAW cells respond poorly to TLR7 and TLR9 ligands but not TLR4 ligands. RAW cells or *Unc93b1* shRNA RAW cells were stimulated with the indicated ligands and TNF- α production was measured by ELISA (representative of two experiments). LPS, lipopolysaccharide; Unst., unstimulated. **c**, *UNC93B1* is required for TLR9 cleavage. Lysates from control RAW-TLR9 cells (vector) or RAW-TLR9 cells transduced with the *Unc93b1* shRNA retrovirus were probed with anti-HA antibodies or anti-tubulin antibodies. **d**, Lysates or purified phagosomes from the same cells described in **c** were analysed for the presence of TLR9 by anti-HA immunoblot. **e**, Overexpression of *UNC93B1* in MEFs enhances TLR9 cleavage. TLR9 in lysates of MEFs stably expressing TLR9-HA or TLR9-HA with *UNC93B1* was detected by anti-HA immunoblot. RAW-TLR9 cells were included as a control for TLR9 cleavage product. **f**, Cells described in **e** were stimulated with 5 μ M CpG or 25 ng ml⁻¹ recombinant mouse IL-1 β or left unstimulated (US). CXCL1/KC (KC) production was measured by ELISA (representative of two experiments). Unless noted otherwise, all data presented are representative of at least three experiments. Error bars represent standard error.

in a 20-fold accumulation of the high-molecular-mass TLR9 species, further supporting the idea that this band is the precursor to the cleavage product.

Two recent reports have shown that inhibitors of cathepsin K (CTSK) and cathepsin B (CTSB) block TLR9 signal transduction in dendritic cells and macrophages, respectively^{13,14}. To address whether specific cathepsins may be involved in TLR9 processing, we tested whether recombinant cathepsins could cleave TLR9 *in vitro*. We observed that recombinant CTSK and CTSS could cleave full-length TLR9 *in vitro* to generate an approximately 80-kDa cleavage product (Fig. 4b). Although these experiments implicate cathepsin family members in TLR9 processing, we were unable to block processing using cathepsin inhibitors. Pulse-chase analysis showed that the cleaved form of TLR9 was still present in RAW-TLR9 cells treated with the broad spectrum inhibitors E64d, pepstatin A and leupeptin as well as with specific inhibitors of CTSK, CTSB and CTSS (Supplementary Fig. 4). With the exception of a CTSB inhibitor, these compounds did not affect TLR9 signalling (Supplementary Fig. 4 and data not shown). Finally, we observed no defect in TLR9 signalling in macrophages and dendritic cells from CTSS-, CTSS/CTSL-, CTSK- and CTSB-deficient mice (Supplementary Fig. 5). Notably, these results disagree with recent work implicating CTSK and CTSB in TLR9 function^{13,14}, although these studies relied heavily on protease inhibitors that may have off-target effects. Instead, our data suggest that none of these proteases is solely responsible for cleaving TLR9 and are consistent with an interpretation that multiple proteases are capable of processing the receptor.

The results described above strongly implicate proteolysis as a step required for TLR9 function, but they do not directly demonstrate that processed TLR9 is a functional receptor. To address this possibility, we tested whether the cleaved form of TLR9 can bind stimulatory CpG oligonucleotides. We found that TLR9 purified from

phagosomes and incubated with biotinylated-CpG oligonucleotides could be efficiently precipitated using streptavidin beads (Fig. 4c). Thus, this processed form of TLR9, which is exclusively in the phagosome, is capable of recognizing ligand. We next compared CpG binding between the full-length and cleaved forms of TLR9. The truncated receptor was pulled down more efficiently than the full-length receptor, suggesting that it may have a higher affinity for ligand (Fig. 4d). This interaction was specific, as binding was disrupted by addition of non-biotinylated CpG oligonucleotide (Fig. 4d).

Because both full-length and processed TLR9 can bind ligand, we sought to determine which form is responsible for signal transduction. The adaptor MyD88 associates with TLR9 after activation, so we examined which form of the receptor recruits MyD88 in response to CpG stimulation. Remarkably, immunoprecipitation of MyD88 in activated cells pulled down only the processed form of TLR9 (Fig. 4e). We were unable to detect any association between MyD88 and the full-length receptor, demonstrating that the processed form of TLR9 is the functional form of the receptor. We next hypothesized that requiring proteolysis may reduce the likelihood that functional receptors will leak to the cell surface and encounter self nucleic acids. To test this possibility, we used a recently reported strategy of fusing *UNC93B1* to a yeast protein domain, Ist2, that directs transport of both *UNC93B1* and TLR9 to the plasma membrane¹⁰. Whereas *UNC93B1*-Ist2 localized to the plasma membrane (Supplementary Fig. 6a), we did not observe significantly altered TLR9 processing or signalling in *UNC93B1*-Ist2 expressing cells (Fig. 4f and Supplementary Fig. 6b). These results suggest that a significant portion of TLR9 may avoid relocalization in *UNC93B1*-Ist2 cells and signal from internal compartments. Thus, we fused TLR9 directly to the Ist2 secretion sequence and examined the signalling and localization of this fusion protein. Consistent with relocalization of the receptor to the cell surface, EndoH analysis indicated that the TLR9-Ist2 fusion protein passed through the Golgi but remained unprocessed (Fig. 4g). Remarkably, TLR9-Ist2 no longer responded to CpG oligonucleotides (Fig. 4h). To ensure that this lack of signalling was not due to the Ist2 domain simply interfering with the TLR9 TIR domain, we generated a TLR9-Ist2 fusion protein in which the key amino acids necessary for plasma membrane targeting have been deleted (TLR9-Ist2(Δ 18))¹⁵. This receptor was processed (Fig. 4g) and regained the ability to respond to CpG oligonucleotides (Fig. 4h), despite the presence of the non-functional Ist2 domain at its C terminus.

Our results reveal that compartmentalized proteolysis of TLR9 represents a key step in nucleic acid recognition. Although both full-length and processed forms of the receptor can bind ligand, only the cleaved form of TLR9 is competent for signal transduction. It remains unclear why processing is necessary for TLR9 activation. In a previous study we showed that a surface-localized chimaeric receptor TLR9N4C (consisting of the TLR9 ectodomain and TLR4 transmembrane and cytosolic domains) was hyper-responsive to self DNA⁴. Unlike TLR9-Ist2, this receptor signals from the cell surface even when fused to Ist2 but is not processed (Fig. 4h right panel and ref. 4). Swapping the transmembrane and cytosolic domains of TLR9 for those of TLR4 has apparently released this receptor from the processing requirement. Recent work suggests that TLR9 undergoes a membrane-proximal conformational change on ligand binding and activation¹⁶. One possibility is that TLR9 cleavage is a prerequisite for this conformational change. This regulatory mechanism may not be required for cell surface receptors like TLR4, explaining the ability of full-length TLR9N4C to signal. Analyses comparing the structure and conformational changes of the full-length and truncated TLR9 receptors will be necessary to resolve these possibilities.

The proteolytic regulatory step described here is consistent with a model in which TLRs involved in nucleic acid sensing are translated as 'pro-receptors' in the ER and only function after being processed in the endolysosomal compartment (Fig. 4i). This scenario is

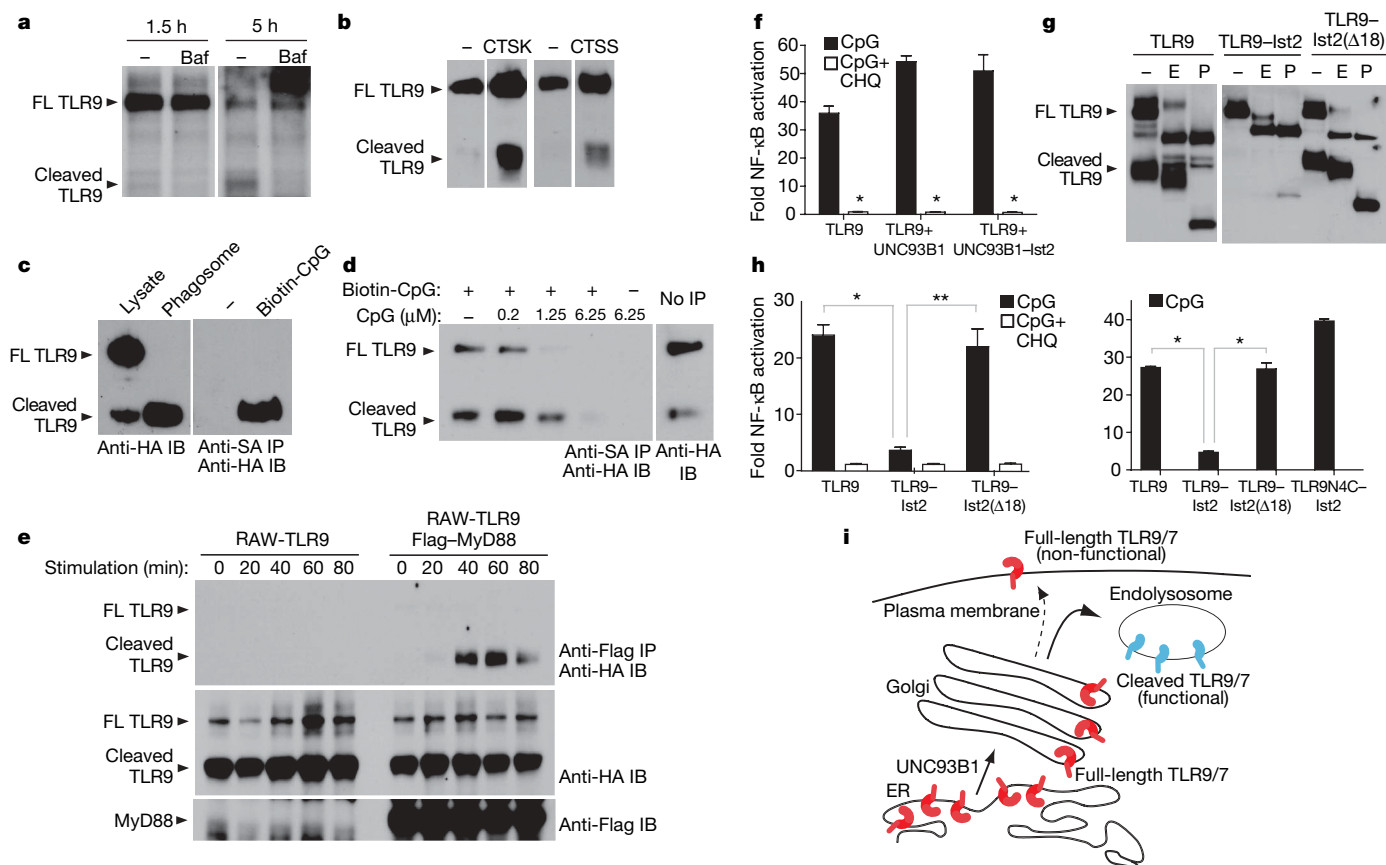


Figure 4 | Processing of TLR9 is necessary to generate a functional receptor. **a**, TLR9 cleavage requires the activity of acid-dependent proteases. Pulse-chase analysis was performed on RAW-TLR9 cells treated with 50 nM bafilomycin A1 (Baf) or DMSO (-) vehicle control. **b**, *In vitro* proteolysis assay performed on TLR9-HA immunoprecipitated from 293T cells. TLR9-HA was incubated with recombinant CTSK, CTSS or no cathepsin (-) and detected by anti-HA immunoblot. **c**, The cleaved form of TLR9 can bind ligand. Left: whole-cell lysates or lysates of purified phagosomes were separated by SDS-PAGE and immunoblotted with anti-HA antibody. Right: lysates of purified phagosomes were incubated with (+) or without (-) 0.2 μM biotin-CpG followed by streptavidin precipitation. TLR9-HA was detected by anti-HA immunoblot. Data are representative of two experiments. **d**, Cell lysates were incubated with (+) or without (-) biotin-CpG and precipitated with streptavidin (SA) as described in **c**. Where indicated, unconjugated CpG was also added, before precipitation. Unmanipulated lysate (No IP) is shown as a control. **e**, Cleaved TLR9

recruits MyD88. RAW cells expressing TLR9-HA and Flag-MyD88 were stimulated with CpG oligonucleotides followed by lysis at the indicated time points. After lysis, MyD88 was immunoprecipitated and TLR9 or MyD88 were detected by anti-HA or anti-Flag immunoblot, respectively. **f**, TLR9 does not signal from the cell surface in UNC93B1-Ist2 cells. 293T cells were transiently transfected with an NF-κB luciferase reporter together with the indicated expression plasmids. Luciferase production was assayed 12 h after stimulation with 3 μM CpG, with or without pre-treatment with 50 μM chloroquine (CHQ). **g**, TLR9, TLR9-Ist2 and TLR9-Ist2(Δ18) expressed in RAW cells were analysed for sensitivity to EndoH or PNGaseF as described in Fig. 2a. **h**, 293T cells were transiently transfected with the indicated expression plasmids and NF-κB activation was measured as described in **f**. **i**, Schematic of compartmentalized proteolysis and regulation of TLR9 and TLR7. For the indicated statistical comparisons, asterisk indicates $P < 0.05$; double asterisk indicates $P < 0.1$. Unless noted otherwise, all data presented are representative of at least three independent experiments.

conceptually analogous to the production of proteases as pro-enzymes that must be processed before they gain enzymatic activity. In this way, compartmentalized processing of TLR9 may regulate where within the cell TLR9 activation can occur. We and others have argued that restricting activation of TLR9 and TLR7 to intracellular compartments may limit responses to self nucleic acid while preserving the ability to recognize foreign nucleic acid^{4,17,18}. Thus, the requirement for processing of TLR9 and TLR7 may represent an evolutionary strategy to ensure proper self/non-self discrimination based on nucleic acid recognition.

METHODS SUMMARY

Detailed descriptions of reagents and experiments can be found in the full Methods. All stable cell lines were generated by retroviral gene transfer. For phagosome isolation, RAW cells were fed latex beads (Polysciences), the cells were disrupted by dounce homogenization, and latex-bead-containing phagosomes were purified on a 62–10% sucrose step gradient in a Beckman SW40Ti centrifuge. After purification, phagosomes were lysed by addition of Triton X-100, and proteins in phagosomes were separated by SDS-polyacrylamide gel electrophoresis (PAGE) and visualized by immunoblot. Pulse-chase analyses

were performed by starving cells of cysteine/methionine for 1 h, pulsing with ³⁵S-cysteine/methionine for 45 min, and chasing in molar excess of unlabelled cysteine/methionine for the indicated time periods. After lysis, radiolabelled proteins were immunoprecipitated and visualized by SDS-PAGE. Deglycosylation assays were performed on immunoprecipitated proteins using endoglycosidase H or PNGaseF (both from NEB) with supplied buffers according to manufacturer's instructions. Proteins were separated by SDS-PAGE and visualized by immunoblot. *In vitro* proteolysis with recombinant cathepsins (Biomol) was performed according to manufacturer guidelines. For MyD88/TLR9 co-immunoprecipitations, cells were lysed in RIPA buffer before or after stimulation with CpG, and protein complexes were precipitated with anti-Flag matrix. For CpG binding assays, TLR9-HA-containing lysates were incubated with biotin-CpG followed by precipitation with streptavidin matrix. In both cases, precipitated proteins were visualized by SDS-PAGE followed by immunoblot. NF-κB luciferase assays were performed by transient transfection of HEK293T cells with an NF-κB luciferase reporter plasmid and the indicated expression plasmids using LTX transfection reagent (Invitrogen) according to manufacturer's instructions. C57/BL/6 mice were purchased from Jackson Laboratories. All knockout mice have been previously described^{19–22}. Mice were housed within animal facilities at UC Berkeley or UCSF according to IACUC guidelines.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 August; accepted 11 September 2008.

Published online 28 September 2008.

1. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature Immunol.* **2**, 675–680 (2001).
2. Iwasaki, A. & Medzhitov, R. Toll-like receptor control of the adaptive immune responses. *Nature Immunol.* **5**, 987–995 (2004).
3. Marshak-Rothstein, A. Toll-like receptors in systemic autoimmune disease. *Nature Rev. Immunol.* **6**, 823–835 (2006).
4. Barton, G. M., Kagan, J. C. & Medzhitov, R. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nature Immunol.* **7**, 49–56 (2006).
5. Latz, E. et al. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nature Immunol.* **5**, 190–198 (2004).
6. Leifer, C. A. et al. TLR9 is localized in the endoplasmic reticulum prior to stimulation. *J. Immunol.* **173**, 1179–1183 (2004).
7. Bell, J. K., Askins, J., Hall, P. R., Davies, D. R. & Segal, D. M. The dsRNA binding site of human Toll-like receptor 3. *Proc. Natl Acad. Sci. USA* **103**, 8792–8797 (2006).
8. Liu, L. et al. Structural basis of Toll-like receptor 3 signaling with double-stranded RNA. *Science* **320**, 379–381 (2008).
9. Brinkmann, M. M. et al. The interaction between the ER membrane protein UNC93B and TLR3, 7, and 9 is crucial for TLR signaling. *J. Cell Biol.* **177**, 265–275 (2007).
10. Kim, Y. M., Brinkmann, M. M., Paquet, M. E. & Ploegh, H. L. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* **452**, 234–238 (2008).
11. Tabeta, K. et al. The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. *Nature Immunol.* **7**, 156–164 (2006).
12. Hacker, H. et al. CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. *EMBO J.* **17**, 6230–6240 (1998).
13. Asagiri, M. et al. Cathepsin K-dependent Toll-like receptor 9 signaling revealed in experimental arthritis. *Science* **319**, 624–627 (2008).
14. Matsumoto, F. et al. Cathepsins are required for Toll-like receptor 9 responses. *Biochem. Biophys. Res. Commun.* **367**, 693–699 (2008).
15. Juschke, C., Wachter, A., Schwappach, B. & Seedorf, M. SEC18/NSF-independent, protein-sorting pathway from the yeast cortical ER to the plasma membrane. *J. Cell Biol.* **169**, 613–622 (2005).
16. Latz, E. et al. Ligand-induced conformational changes allosterically activate Toll-like receptor 9. *Nature Immunol.* **8**, 772–779 (2007).
17. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis e Sousa, C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**, 1529–1531 (2004).
18. Lund, J. M. et al. Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl Acad. Sci. USA* **101**, 5598–5603 (2004).
19. Deussing, J. et al. Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. *Proc. Natl Acad. Sci. USA* **95**, 4516–4521 (1998).
20. Mallen-St & Clair, J. et al. Cathepsins L and S are not required for activation of dipeptidyl peptidase I (cathepsin C) in mice. *Biol. Chem.* **387**, 1143 (2006).
21. Saftig, P. et al. Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. *Proc. Natl Acad. Sci. USA* **95**, 13453–13458 (1998).
22. Shi, G.-P. et al. Cathepsin S required for normal MHC class II peptide loading and germinal center development. *Immunity* **10**, 197–206 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank members of the Barton and Vance laboratories for discussions, S. Akira for providing TLR knockout mice, and R. Vance and R. Medzhitov for advice. We acknowledge the CHPS Imaging Core and P. Herzmark for assistance with microscopy. This work was supported by a Novel Research Grant from The Lupus Research Institute (G.M.B.), the Hellman Faculty Fund (G.M.B.), and NIH grants HL67204 (H.A.C.), CA009179 (S.E.E.) and AI072429 (G.M.B.).

Author Contributions S.E.E. and G.M.B. planned experiments and wrote the manuscript. S.E.E., B.L.L., L.L. and K.E.W. performed the experiments. G.-P.S. and H.A.C. contributed reagents. All authors read and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.M.B. (barton@berkeley.edu).

METHODS

Reagents. All chemicals and reagents, unless noted otherwise, were purchased from Fisher Scientific. Anti-HA (clone 3F10) and anti-Flag (M2) affinity matrices were purchased from Roche and Sigma, respectively. Latex beads were purchased from Polysciences. CpG oligonucleotides (TCCATGACG-TTCTGACGTT) with phosphorothioate linkages and 5'-biotinylated CpG oligonucleotides (of the same sequence) were purchased from Invitrogen and IDT. LPS, R848, BLP and biotinylated LPS were purchased from InvivoGen. All antibodies for flow cytometry and ELISAs were purchased from eBioscience. The following antibodies were used for immunoblotting: anti-HA (Roche), anti-Flag M2 (Sigma), anti-calnexin (Stressgen), anti-HSP90B (Thermo Scientific), anti-LAMP1 (Santa Cruz), anti-tubulin (EMD), anti-nuclear pore complex (Covance), goat anti-rat-HRP and sheep anti-mouse HRP (GE Amersham).

Mice. C57BL/6 and 129Sv mice were purchased from Jackson Laboratories. *Tlr2*^{-/-} *Tlr4*^{-/-} *Tlr9*^{-/-} mice were generated by intercrossing individual knockout mice, provided by S. Akira. *Ctsk*^{-/-} (ref. 21), *Ctss*^{-/-} (ref. 22) and *Ctss*^{-/-} *Ctsl*^{-/-} (ref. 20) mice (all backcrossed to C57BL/6) and *CtsB*^{-/-} (ref. 19) mice (on the 129Sv background) have been previously described. All mice were housed within the animal facilities at the University of California at Berkeley or University of California at San Francisco according to IACUC guidelines.

Cell lines, plasmids and tissue culture. RAW264 cells were purchased from ATCC and cultured in RPMI supplemented with 10% FCS, L-glutamine, penicillin/streptomycin, sodium pyruvate and HEPES (all from Invitrogen). Unless otherwise noted, stable lines were generated by transducing cells with MSCV2.2 retroviruses encoding the target cDNA. For RAW-TLR9, RAW-TLR7 and RAW-TLR4 cells, cDNA encoding the HA epitope was inserted at the 3' end of each TLR. Flag-TLR9 and Flag-MyD88 have been previously described²³. UNC93B-Ist2 was constructed as previously described¹⁰. TLR9-Ist2 and TLR9N4C-Ist2 were made by fusing the 69-residue Ist2 tag to the 3' end of TLR9-HA or TLR9N4C-HA, respectively.

Bone-marrow-derived conventional dendritic cells, macrophages and plasmacytoid dendritic cells were differentiated, as previously described⁴, in RPMI supplemented with GM-CSF containing supernatant, M-CSF containing supernatant, or 25 ng ml⁻¹ recombinant Flt3L (R&D), respectively. To generate conventional dendritic cells and plasmacytoid dendritic cells expressing TLR9-HA, haematopoietic stem cells (HSCs) from *Tlr2*^{-/-} *Tlr4*^{-/-} *Tlr9*^{-/-} mice were transduced with retroviruses encoding TLR9-HA as previously described⁴. Briefly, mice were injected with 5-fluorouracil (Sigma) to enrich for HSCs. HSCs were cultured in IL-6, IL-3 and SCF (all from R&D) over 4 days and transduced with retroviruses. After 5 days, HSCs were differentiated into conventional dendritic cells or plasmacytoid dendritic cells as described above.

Splenic dendritic cells were isolated from C57BL/6 or *Ctsk*^{-/-} mice by magnetic cell sorting using anti-CD11c conjugated magnetic beads (Miltenyi Biotech) and an AutoMACS cell sorter (Miltenyi Biotech). Enrichment, as measured by FACS, was greater than 70%.

Immunoprecipitation and western blot analyses. Cells were lysed in RIPA buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with Complete protease inhibitor cocktail (Roche). After incubation on ice, lysates were cleared of insoluble material by centrifugation. For immunoprecipitations, lysates were incubated with anti-HA matrix or anti-Flag matrix and precipitated proteins were separated by SDS-PAGE and transferred to Immobilon PVDF membrane (Millipore). Membranes were probed with the indicated antibodies and developed by ECL chemiluminescence (Pierce). Relative band intensities were quantified by counting pixels per band using ImageJ software.

Streptavidin pull downs were performed on cells or purified phagosomes lysed in TNT buffer (20 mM Tris pH 8.0, 200 mM NaCl, 1% Triton X-100) and cleared of insoluble debris. Lysates were incubated for 1 h with 0.2 μM biotin-CpG and/or unlabelled CpG on ice, added to streptavidin-agarose (Invitrogen, pre-blocked with 1% BSA-PBS), rotated 1 h at 4 °C, and washed four times in lysis buffer. Precipitates were boiled in SDS buffer, separated by SDS-PAGE, and probed by anti-HA immunoblot.

Microscopy. Co-localization studies were performed on a Zeiss 510 Axioplan META microscope. Coverslips were coated with poly-D-lysine (Sigma), washed extensively with water and allowed to dry (1 h). Cells were allowed to settle overnight. The next day, coverslips were washed with PBS, fixed with 4%

PFA/PBS, and permeabilized with 0.5% Triton X-100/PBS. After washing with 0.1% Triton X-100/PBS, cells were blocked (5% goat sera, 2% BSA, 0.1% sodium azide, 0.1% Triton X-100/PBS). Slides were stained in blocking buffer with anti-Flag, anti-HA, anti-calnexin or anti-LAMP1 (see antibodies used for western blot). Secondary antibodies were donkey anti-rabbit Cy5, donkey anti-rabbit Cy3, goat anti-rat Cy3, goat anti-mouse Cy3 (Jackson) or donkey anti-rat Alexa680 (Invitrogen).

Visualization of UNC93B1-GFP and UNC93B1-GFP-Ist2 was performed on a Nikon E800 fluorescent microscope. Cells were plated and fixed as described above.

Luciferase assays. Transient transfections were performed in OptiMem (Invitrogen) with LTX transfection reagent (Invitrogen) according to manufacturer's guidelines. 293T cells were transiently co-transfected with the indicated expression plasmid(s) and a plasmid encoding an NF-κB luciferase reporter (with five tandem NF-κB binding sites, provided by R. Vance). Cells were stimulated with 3 μM CpG oligonucleotide after 24 h and lysed by passive lysis after an additional 12 h. Luciferase activity was measured on a LMaxII-384 luminometer (Molecular Devices).

Pulse-chase. Cells were starved for 1 h in cysteine/methionine-free media, then pulsed with 0.1 mCi ³⁵S-cysteine/methionine (Perkin-Elmer). After a 45-min pulse, cells were washed and cultured in 5 ml chase media with 10,000-fold molar excess of L-cysteine, L-methionine or harvested as the zero time point. Time points were harvested as follows: cells were washed twice in 2 ml PBS and lysed in 1 ml RIPA plus protease inhibitor cocktail. For protease inhibitor experiments, inhibitors were added in chase media and supplemented every 2 h at the following concentrations: 50 μM E64d (EMD), 100 μM pepstatin A (Sigma), 100 μM leupeptin (Sigma), 10 nM cathepsin K inhibitor III (Cbz-Leu-NH-CH₂-CO-CH₂-NH-Leu-Cbz, Sigma), 30 μM CtsB inhibitor (Me047, EMD), 2 μM CtsS inhibitor (-Phe-Leu-COCHO·H₂O, EMD). Baflomycin A1 (Sigma) was used at 50 nM and brefeldin A (Sigma) was used at 10 μg ml⁻¹.

Phagosome isolation. 4 × 10⁷ RAW cells were incubated with ~10⁸ 2 μM latex beads (Polysciences) for various time points. After rigorous washing in PBS, RAW cells were scraped into 3 ml sucrose homogenization buffer (SHB: 250 μM sucrose, 3 mM imidazole pH 7.4) and pelleted by centrifugation. Cells were resuspended in 1 ml SHB plus protease inhibitor cocktail (Roche) and EDTA and disrupted by 25 strokes in a steel dounce homogenizer. The disrupted cells were gently rocked for 10 min on ice and centrifuged at 8,000g to remove intact cells, nuclei and gross debris. The remaining supernatant was further processed to separate intact phagosomes. A total of 40 μl was set aside as the 'lysate' sample and lysed by addition of 1% Triton X-100. The remaining supernatant was mixed with an equal volume of 60% sucrose, and applied to a sucrose step gradient as follows: 1 ml 60% sucrose, 2 ml homogenate, 2 ml 32% sucrose, 2 ml 20% sucrose, 2 ml 10% sucrose. The gradient was spun at 100,000g for 1 h in a SW40 rotor. Following centrifugation, phagosomes were harvested from the 20–10% sucrose interface. Phagosomes were diluted in 10 ml PBS and spun at 40,000g for 15 min in an SW40 rotor. The phagosome pellet was resuspended in 40 μl SHB, protease inhibitor cocktail, EDTA and 1% Triton X-100. Protein concentrations were normalized by BCA assay, and lysates were analysed by western blot.

In vitro proteolysis assay. Recombinant cathepsin S and recombinant pro-cathepsin K were purchased from BioMol. 293T cells stably expressing TLR9-HA were lysed in RIPA buffer and TLR9 was immunoprecipitated with anti-HA agarose. Immunoprecipitates were washed four times in cathepsin reaction buffer (50 mM potassium phosphate pH 6.5, 50 mM NaCl, 2 mM EDTA, 0.01% Triton X-100, 0.5 mM dithiothreitol) followed by incubation with or without cathepsins. Pro-cathepsin K was activated according to manufacturer's instructions (that is, incubation in 32.5 mM NaOAc pH 3.5 for 25 min at 24 °C). Activated cathepsin K was immediately incubated with immunoprecipitated TLR9-HA in cathepsin reaction buffer for 45 min at 24 °C. Cathepsin S (0.0024 units) was added directly to immunoprecipitated TLR9-HA in cathepsin reaction buffer and incubated for 45 min at 40 °C. Enzymes were inactivated by the addition of SDS loading buffer and boiling. Reactions were resolved by SDS-PAGE and anti-HA immunoblot (as described above).

23. Horng, T., Barton, G. M. & Medzhitov, R. TIRAP: an adapter molecule in the Toll signaling pathway. *Nature Immunol.* 2, 835–841 (2001).

Regulation of *ERBB2* by oestrogen receptor–PAX2 determines response to tamoxifen

Antoni Hurtado¹, Kelly A. Holmes¹, Timothy R. Geistlinger², Iain R. Hutcheson³, Robert I. Nicholson³, Myles Brown², Jie Jiang⁴, William J. Howat¹, Simak Ali⁴ & Jason S. Carroll¹

Crosstalk between the oestrogen receptor (ER) and *ERBB2*/HER-2 pathways has long been implicated in breast cancer aetiology and drug response¹, yet no direct connection at a transcriptional level has been shown. Here we show that oestrogen–ER and tamoxifen–ER complexes directly repress *ERBB2* transcription by means of a *cis*-regulatory element within the *ERBB2* gene in human cell lines. We implicate the paired box 2 gene product (PAX2), in a previously unrecognized role, as a crucial mediator of ER repression of *ERBB2* by the anti-cancer drug tamoxifen. We show that PAX2 and the ER co-activator AIB-1/SRC-3 compete for binding and regulation of *ERBB2* transcription, the outcome of which determines tamoxifen response in breast cancer cells. The repression of *ERBB2* by ER–PAX2 links these two breast cancer subtypes and suggests that aggressive *ERBB2*-positive tumours can originate from ER-positive luminal tumours by circumventing this repressive mechanism. These data provide mechanistic insight into the molecular basis of endocrine resistance in breast cancer.

The genomic mapping of ER-binding sites has provided insight into how ER functions in breast cancer cells, including the finding that ER rarely binds to promoter regions and that loading of ER on the chromatin requires the presence of pioneer factors, such as FoxA1 (refs 2–4). We have replicated genome-wide ER chromatin immunoprecipitation (ChIP)-on-chip analyses in ER-positive MCF-7 cells. Identification of the ER-binding sites with a false discovery rate of 5% revealed 8,525 ER sites, with high representation (86%) of the published ER binding profile² (Supplementary Data 2). Included within the new, more extensive list was an ER-binding site within the intron of the *ERBB2*/HER-2 genomic region (Fig. 1a). Sequence analysis of all 8,525 ER-binding sites revealed a statistical enrichment ($P < 0.0001$) for the PAX transcription factor motif (GTCANGN(A/G)T) (Fig. 1b). Little is known about the function of PAX proteins in hormone signalling; however, PAX2 was shown to be expressed in a subset of breast cancers and was recently identified as a tamoxifen-regulated effector in endometrial cancer cells^{5,6}.

Tamoxifen is one of the most successful and effective therapies in the treatment of breast cancer, but resistance to tamoxifen is common⁷. Tamoxifen-resistant breast tumours are characterized by elevated *ERBB2* levels⁸, and ER-positive cell line models overexpressing *ERBB2* acquire resistance to tamoxifen⁹. We assessed the binding of PAX2 to selected ER-binding sites adjacent to important oestrogen-regulated genes, including the newly identified binding site within the *ERBB2* gene. PAX2 was generally recruited only after treatment with tamoxifen, with the exception of the ER-binding site within *ERBB2* (Fig. 1c), where PAX2 was recruited to the ER-binding site after both treatment with oestrogen and treatment with tamoxifen. Given previous evidence that *ERBB2* could be repressed by both oestrogen¹⁰

and tamoxifen¹¹, we proposed that PAX2 might be functioning as a general ER-associated transcriptional repressor and that the ER-binding site within *ERBB2* might be a *cis*-regulatory element for active repression by ER. Indeed, analysis confirmed that levels of *ERBB2* messenger RNA are decreased by oestrogen and by tamoxifen in our MCF-7 cells (Fig. 1d). Co-immunoprecipitation experiments showed that ER and PAX2 form a complex after treatment with tamoxifen (Supplementary Data 3) and re-ChIP experiments (ChIP followed by release of chromatin and re-ChIP using an antibody against a different protein) confirmed that ER and PAX2 occupy the same ER-binding site within the *ERBB2* gene simultaneously, after treatment with tamoxifen (Supplementary Data 3). Furthermore, we experimentally verified this ER-binding site as the *cis*-regulatory element for the *ERBB2* gene (Supplementary Data 4). This *cis*-regulatory region is independent of a previously identified regulatory region¹⁰, although this previously characterized region might have an indirect function in the regulation of *ERBB2* transcription.

ER-positive luminal tumours with the poorest prognosis tend to have elevated *ERBB2* levels¹² and up to half of *ERBB2*-positive tumours are also positive for ER¹³. We therefore proposed that the anti-proliferative effects of tamoxifen treatment require repression of *ERBB2*, and that breast cancers can potentially acquire tamoxifen resistance by amplifying the *ERBB2* locus or by deregulating the control mechanisms that normally repress *ERBB2* transcription. Unlike tamoxifen, repression of *ERBB2* by oestrogen may not be a critical event, because cell proliferation by oestrogen probably results from the oestrogen-mediated upregulation of numerous oncogenes.

To investigate the possible role for PAX2 in the oestrogen-mediated and tamoxifen-mediated repression of *ERBB2*, we specifically silenced PAX2 with short interfering RNA (siRNA). Immunoblotting revealed efficient knockdown of PAX2 protein levels, but no significant effect on ER levels (Fig. 2a). In control transfected cells, oestrogen and tamoxifen both rapidly repressed *ERBB2* mRNA (Fig. 2b), but PAX2 siRNA abrogated this inhibition and consequently *ERBB2* transcription and *ERBB2* protein levels were elevated in the presence of both treatment with oestrogen and treatment with tamoxifen (Figs 2a and 2b). This coincided with an accumulation of phosphorylated RNA polymerase II (PolII) at the promoter (the longer isoform) of *ERBB2* after treatment with oestrogen and with tamoxifen in the presence of PAX2 siRNA (Supplementary Data 5). Relative to the control, treatment of PAX2 siRNA-transfected cells with tamoxifen resulted in an increase in cell number (Fig. 2c), reversing the growth arrest observed after treatment with tamoxifen. These experiments were reproduced with an additional PAX2 siRNA (Supplementary Data 6). Pretreatment of cells with an anti-*ERBB2* antibody (Herceptin) blocked the PAX2

¹Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. ²Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA. ³Tenovus Centre for Cancer Research, Welsh School of Pharmacy, Cardiff University, Cardiff CF10 3XF, UK. ⁴Imperial College London, Hammersmith Hospital, London W12 0NN, UK.

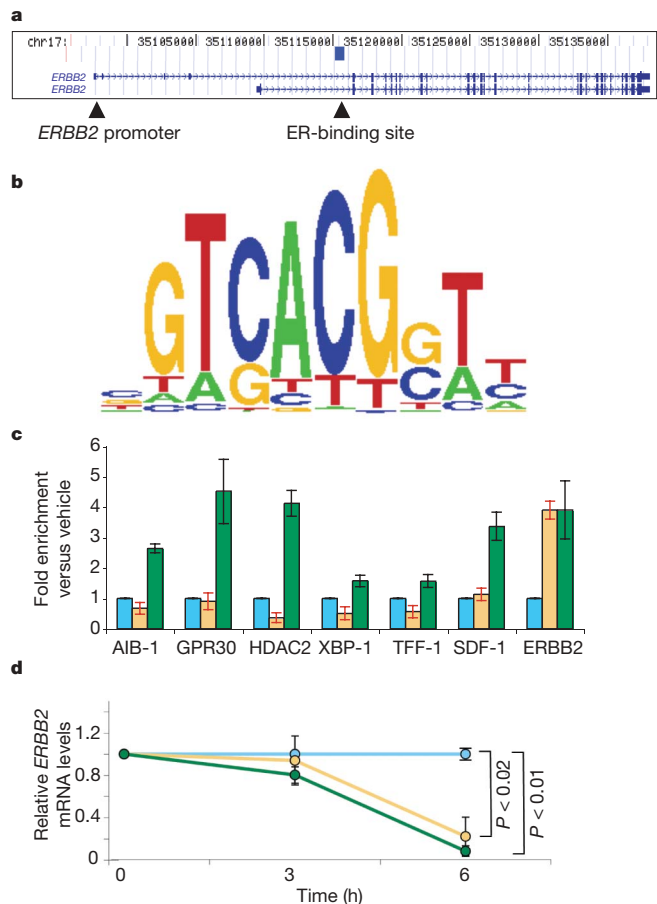


Figure 1 | ER ChIP-on-chip reveals a binding site within the *ERBB2* gene region that is bound by PAX2. **a**, Schematic representation of the *ERBB2* gene locus and the intronic ER-binding site, as defined by ER ChIP-on-chip experiments, on chromosome 17. Both isoforms of *ERBB2* are shown. **b**, PAX motif enriched within ER-binding sites. **c**, PAX2 ChIP after treatment with vehicle (blue bars), oestrogen (yellow bars) or tamoxifen (green bars). **d**, Changes in *ERBB2* mRNA by real-time RT-PCR after treatment with vehicle (blue), oestrogen (yellow) or tamoxifen (green). All graphical results are shown as means and s.d. for three independent replicates.

siRNA-mediated cell growth, confirming that the increased cell growth after PAX2 silencing was due primarily to the increase in *ERBB2* levels (Fig. 2c).

We recapitulated these findings in another ER-positive breast cancer cell line (ZR75-1 cells). Tamoxifen repressed *ERBB2* mRNA and *ERBB2* protein levels in ZR75-1 cells, and this repression was inhibited after silencing of PAX2. Similarly to MCF-7 cells, PAX2 siRNA reversed the growth inhibitory effects of tamoxifen, such that ZR75-1 cells acquired tamoxifen resistance in the absence of PAX2 (Supplementary Data 8).

As well as having elevated *ERBB2* levels, tamoxifen-resistant breast cancers are also characterized by increased levels of the ER co-activator AIB-1 (amplified in breast cancer-1; also known as SRC-3 (ref. 8)). AIB-1 promotes tumorigenesis^{14,15} and is essential for *ERBB2*-driven oncogenesis in mice¹⁶. We assessed whether AIB-1 could compete with PAX2 for binding to the *ERBB2* cis-regulatory element, an event that may contribute to the elevated *ERBB2* levels associated with tamoxifen-resistant tumours⁸. ChIP showed decreased AIB-1 binding at the *ERBB2* cis-regulatory element after treatment with oestrogen and treatment with tamoxifen (Fig. 2d), probably as a result of displacement by PAX2. This was proved by inhibiting PAX2 with siRNA, which consequently allowed oestrogen-mediated and tamoxifen-mediated recruitment of AIB-1 to the *ERBB2* cis-regulatory element (Fig. 2d).

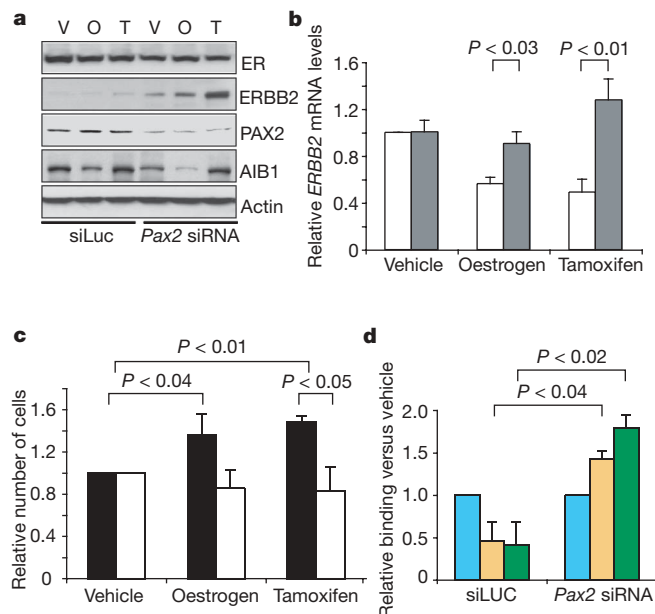


Figure 2 | Specific silencing of PAX2 reverses the tamoxifen-mediated repression of *ERBB2* and growth arrest. **a**, siRNA to PAX2 was transfected into hormone-depleted MCF-7 cells and stimulated with vehicle (V), oestrogen (O) or tamoxifen (T), and total protein was immunoblotted. **b**, Control siRNA (siLuc; white bars) or PAX2 siRNA (grey bars) was transfected and *ERBB2* mRNA levels were assessed. **c**, PAX2 siRNA was transfected into cells in the presence of control (black bars) or Herceptin (anti-*ERBB2* antibody; white bars), after which cells were collected and the total number of viable cells was determined. The control transfection (siLuc) data are in Supplementary Fig. 7. **d**, Transfection of control siRNA (siLuc) or PAX2 siRNA was performed as described and cells were treated with vehicle (blue bars), oestrogen (yellow bars) or tamoxifen (green bars). Binding of AIB-1 to the *ERBB2* enhancer was determined by ChIP. All graphical results are shown as means and s.d. for three independent replicates.

We subsequently showed that expression of AIB-1 competes with PAX2 for binding to the *ERBB2* cis-regulatory element and that this results in an increase in *ERBB2* transcription and an increase in cell proliferation in the presence of tamoxifen (Supplementary Data 9). Elevated AIB-1 levels block PAX2 binding and *ERBB2* gene repression, thereby reversing the antiproliferative effects of tamoxifen. This suggests that a stoichiometric balance between the co-activator AIB-1 and the putative repressor PAX2 impinges on the binding and regulation of *ERBB2*, providing mechanistic insight into the function of AIB-1 in the tamoxifen response¹⁷. MCF-7 cells already have elevated AIB-1 levels as a result of a genomic amplification of the *AIB-1* locus¹⁴, but they also have increased PAX2 protein levels (data not shown), potentially explaining why they retain sensitivity to treatment with tamoxifen. However, we were also able to show that AIB-1 expression could reverse the anti-proliferative effects of tamoxifen in T47D cells, a cell line that does not already have elevated AIB-1 levels¹⁴ (Supplementary Data 9). These data confirm a general role for AIB-1 in reversing tamoxifen responsiveness in ER-positive breast cancer cell lines.

The role of *ERBB2* in tamoxifen resistance is demonstrated by data showing that tamoxifen-resistant breast cancer cell lines can be inhibited by treatment with anti-*ERBB2* antibodies (Herceptin)¹⁸. We investigated the hypothesis that PAX2 is required for repression of *ERBB2* and that tamoxifen resistance may be due to alterations in this pathway. We used an MCF-7 cell line model that had been grown in the presence of tamoxifen and had acquired resistance¹⁸. These tamoxifen-resistant (Tam-R) cells have elevated *ERBB2* levels but do not have amplification of the *ERBB2* locus¹⁸. In wild-type MCF-7 cells, tamoxifen repressed *ERBB2* mRNA levels (Fig. 2b) and *ERBB2* protein levels (by 40%) (Fig. 3a) as expected, but *ERBB2* levels

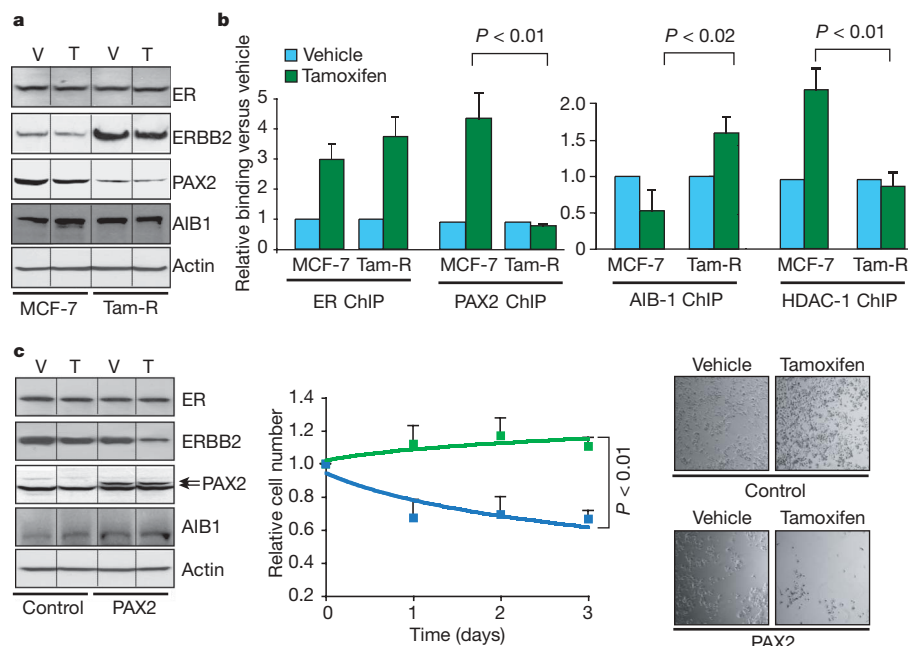


Figure 3 | PAX2 restores sensitivity to tamoxifen in tamoxifen-resistant breast cancer cells. **a**, Total protein from wild-type or tamoxifen-resistant (Tam-R) MCF-7 cells was immunoblotted. V, vehicle; T, tamoxifen. **b**, ChIP of ER, PAX2, AIB-1 and HDAC-1 at the ER-binding site in the *ERBB2* gene in wild-type and Tam-R cells after treatment with vehicle (blue bars) or tamoxifen (green bars). **c**, Control or PAX2-expressing plasmids were transfected into Tam-R cells, followed by treatment with vehicle (V) or tamoxifen (T). Total protein was immunoblotted. After PAX2 expression in Tam-R cells, total numbers of viable cells were determined after treatment with vehicle or tamoxifen. The data are tamoxifen treatment versus vehicle. Green, control; blue, PAX2. The immunoblots have been cropped; the original figures are in Supplementary Data 11. All graphical results are shown as means and s.d. for three independent replicates.

were elevated in Tam-R cells and did not decrease in response to treatment with tamoxifen (Fig. 3a). Western blot analysis comparing wild-type and Tam-R MCF-7 cells revealed no changes in ER protein levels, supporting clinical studies showing that changes in ER levels are not a general mechanism for tamoxifen-resistant breast cancers^{19,20}. AIB-1 protein levels were also unaltered, but PAX2 protein levels were lower in Tam-R cells (Fig. 3a), providing a potential explanation for the elevated ERBB2 levels in these tamoxifen-resistant cells.

Tamoxifen-mediated ER recruitment to the *ERBB2* cis-regulatory element was assessed in the Tam-R cells and was shown to be similar to that in wild-type MCF-7 cells (Fig. 3b). However, as suspected from the lower PAX2 levels in Tam-R cells, PAX2 binding was significantly decreased in the Tam-R cells. Similarly, binding of histone deacetylase 1 (HDAC-1) was shown to occur only in the wild-type cells and not in the Tam-R cells, confirming that active repression occurs at the *ERBB2* cis-regulatory element in wild-type cells but not in the Tam-R cells. In contrast, tamoxifen-mediated AIB-1 recruitment was elevated in Tam-R cells (Fig. 3b), despite unaltered AIB-1 levels. To test the hypothesis that the decreased PAX2 levels contributed to the increased expression of ERBB2 and the altered response to tamoxifen in the Tam-R cells, we reintroduced PAX2 into these cells (Fig. 3c). After overexpression of PAX2, tamoxifen was now able to repress *ERBB2* mRNA (Supplementary Data 12) and ERBB2 protein levels (Fig. 3c) in Tam-R cells. The overexpression of PAX2 resulted in decreased binding of PolII to the *ERBB2* promoter and decreased binding of AIB-1 to the ER-binding site (Supplementary Data 12), strengthening the hypothesis that AIB-1 and PAX2 compete for binding and regulation of the *ERBB2* gene. Active gene repression by tamoxifen was restored by PAX2 expression, as indicated by the

recruitment of HDAC-1 (Supplementary Data 12). PAX2 was subsequently shown to be a critical regulator of cellular proliferation, because expression of PAX2 restored the ability of tamoxifen to inhibit cell growth in these previously resistant cells (Fig. 3c).

We recapitulated these findings in BT-474 breast cancer cells, which are ER positive but resistant to tamoxifen²¹, probably as a result of a genomic amplification of the *ERBB2* locus²². These cells therefore represent another possible mechanism of acquired tamoxifen resistance, whereby amplification of the *ERBB2* locus can overcome the growth inhibitory effects imposed by tamoxifen in ER-positive breast cancers^{23,24}. Expression of PAX2 in BT-474 cells resulted in tamoxifen-mediated repression of *ERBB2* mRNA and ERBB2 protein levels (Supplementary Data 13) and resulted in tamoxifen-dependent inhibition of cell growth (Supplementary Data 13), such that growth inhibitory effects of tamoxifen were restored, even in the presence of the amplified *ERBB2* locus.

Our findings suggest that PAX2 is a key deterministic component in the tamoxifen response. To confirm these findings in primary breast cancer, we performed PAX2 immunohistochemistry on 109 ER-positive breast cancer samples²⁵, all of which had been treated with tamoxifen. Of these 109 tumours, 68 were PAX2 positive and 41 were PAX2 negative. Tumours with positive PAX2 staining corresponded to a significantly improved recurrence-free survival in patients relative to PAX2-negative tumours ($P < 0.0001$) (Supplementary Data 14). Furthermore, within the PAX2-positive tumours only, those that were also positive for AIB-1 had a worse clinical outcome than the tumours that were AIB-1 negative (Fig. 4). The tumours that were PAX2 positive and AIB-1 negative had the best prognosis of all, with a recurrence rate of only 5.8% (Fig. 4). Cox

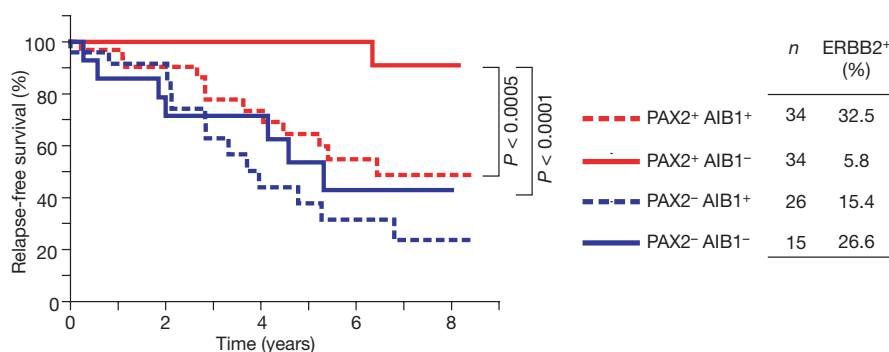


Figure 4 | PAX2 predicts clinical outcome in breast cancer patients. Kaplan-Meier curve representing the percentage relapse-free survival in tumours based on staining for PAX2 and AIB-1 ($n = 109$). The percentage of ERBB2-overexpressing tumours within each category is shown.

regression analysis confirmed an inverse dependent relationship between PAX2 and AIB-1 levels in determining relapse ($P < 0.03$). The PAX2-positive, AIB-1-negative tumours also had the lowest percentage of ERBB2-positive staining (Fig. 4), supporting our hypothesis that a balance between PAX2 and AIB-1 ultimately dictates ERBB2 expression and determines tamoxifen efficacy.

Endocrine resistance is a significant problem in breast cancer treatment. One of the few validated clinical features of tamoxifen-resistant breast cancer is the combined elevation of the AIB-1 and ERBB2 pathways⁸. We now provide evidence that PAX2 is a critical tamoxifen-recruited transcriptional repressor of the *ERBB2* gene and that increased AIB-1 expression can outcompete PAX2 binding, directly resulting in increased ERBB2 expression. Alterations in AIB-1–PAX2 stoichiometry dictate the efficacy of tamoxifen in breast cancers (a schematic model of these events is shown in Supplementary Data 1). These new data suggest an intrinsic transcriptional link between tumours driven by ER and those driven by ERBB2, which together account for a significant majority of all breast cancers. The role of PAX2 as a repressor is unexpected, because PAX2 is generally a transcriptional activator and was shown to be encoded by a tamoxifen-regulated gene that can induce endometrial cancer⁶. Given that tamoxifen has antiproliferative effects in the breast but possesses agonist properties in the endometrium²⁶, it is possible that PAX2 may have tissue specific effects and may be one of the primary determinants for selective oestrogen receptor modulator (SERM) action in female reproductive tissue.

METHODS SUMMARY

The MCF-7, T47D, ZR75-1 and BT-474 human cell lines were grown as described previously²⁷. Tam-R cells were derived by long-term exposure to tamoxifen¹⁸ and grown under the same conditions as wild-type MCF-7 cells. Genome-wide ER ChIP-on-chip experiments were performed in MCF-7 cells in duplicate, as described previously². ER-binding-site analyses were determined with MAT²⁸, with a false discovery rate of 5%. The analysis of motif enrichments was performed with the CEAS program (<http://ceas.cbi.pku.edu.cn/>). For mRNA experiments, cells were deprived of hormones as described previously²⁹. Total RNA was collected and RT-PCR was performed as described previously². ChIP experiments were performed in hormone-depleted medium as described previously³. Proliferation assays were performed in complete medium. siRNA experiments were performed as described previously² in hormone-depleted medium. For immunohistochemistry, 109 ER-positive breast cancer sections were collected and processed as described previously²⁵. Immunohistochemistry for PAX2 was performed on an automated BondMax Immunostainer (Leica) with anti-PAX2 antibody (ab38738; Abcam) and anti-AIB-1 antibody (611105; Transduction Laboratories). Immunohistochemistry for ERBB2 was performed as described previously²⁵. Statistical analysis was performed with two-tailed paired *t*-tests, and $P < 0.05$ was considered statistically significant. In all figures, graphical results are shown as means and s.d. for a minimum of three independent replicates. Kaplan–Meier curve statistics were determined with a log-rank test. PAX2 and AIB-1 relationships were established by Cox regression analysis.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 July; accepted 2 October 2008.

Published online 12 November 2008.

1. Ali, S. & Coombes, R. C. Endocrine-responsive breast cancer and strategies for combating resistance. *Nature Rev.* **2**, 101–112 (2002).
2. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nature Genet.* **38**, 1289–1297 (2006).
3. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
4. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
5. Muratovska, A. *et al.* Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival. *Oncogene* **22**, 7989–7997 (2003).
6. Wu, H. *et al.* Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis. *Nature* **438**, 981–987 (2005).
7. Clarke, R. *et al.* Cellular and molecular pharmacology of antiestrogen action and resistance. *Pharmacol. Rev.* **53**, 25–71 (2001).

8. Osborne, C. K. *et al.* Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J. Natl Cancer Inst.* **95**, 353–361 (2003).
9. Benz, C. C. *et al.* Estrogen-dependent, tamoxifen-resistant tumorigenic growth of MCF-7 cells transfected with HER2/neu. *Breast Cancer Res. Treat.* **24**, 85–95 (1992).
10. Bates, N. P. & Hurst, H. C. An intron 1 enhancer element mediates oestrogen-induced suppression of ERBB2 expression. *Oncogene* **15**, 473–481 (1997).
11. Frasar, J. *et al.* Selective estrogen receptor modulators: discrimination of agonistic versus antagonistic activities by gene expression profiling in breast cancer cells. *Cancer Res.* **64**, 1522–1533 (2004).
12. Kun, Y. *et al.* Classifying the estrogen receptor status of breast cancers by expression profiles reveals a poor prognosis subpopulation exhibiting high expression of the ERBB2 receptor. *Hum. Mol. Genet.* **12**, 3245–3258 (2003).
13. Piccart-Gebhart, M. J. *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N. Engl. J. Med.* **353**, 1659–1672 (2005).
14. Anzick, S. L. *et al.* AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* **277**, 965–968 (1997).
15. Torres-Arzuayus, M. I. *et al.* High tumor incidence and activation of the PI3K/AKT pathway in transgenic mice define AIB1 as an oncogene. *Cancer Cell* **6**, 263–274 (2004).
16. Fereshteh, M. P. *et al.* The nuclear receptor coactivator amplified in breast cancer-1 is required for Neu (ErbB2/HER2) activation, signaling, and mammary tumorigenesis in mice. *Cancer Res.* **68**, 3697–3706 (2008).
17. Louie, M. C. *et al.* ACTR/AIB1 functions as an E2F1 coactivator to promote breast cancer cell proliferation and antiestrogen resistance. *Mol. Cell. Biol.* **24**, 5157–5171 (2004).
18. Knowlden, J. M. *et al.* Elevated levels of epidermal growth factor receptor/c-erbB2 heterodimers mediate an autocrine growth regulatory pathway in tamoxifen-resistant MCF-7 cells. *Endocrinology* **144**, 1032–1044 (2003).
19. Johnston, S. R. *et al.* Changes in estrogen receptor, progesterone receptor, and pS2 expression in tamoxifen-resistant human breast cancer. *Cancer Res.* **55**, 3331–3338 (1995).
20. Green, K. A. & Carroll, J. S. Oestrogen-receptor-mediated transcription and the influence of co-factors and chromatin state. *Nature Rev.* **7**, 713–722 (2007).
21. Zhou, Y. *et al.* The NF- κ B pathway and endocrine-resistant breast cancer. *Endocr. Relat. Cancer* **12** (Suppl 1), S37–S46 (2005).
22. Clark, J. *et al.* Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. *Genes Chromosomes. Cancer* **34**, 104–114 (2002).
23. Dowsett, M. Overexpression of HER-2 as a resistance mechanism to hormonal therapy for breast cancer. *Endocr. Relat. Cancer* **8**, 191–195 (2001).
24. Dowsett, M. *et al.* HER-2 amplification impedes the antiproliferative effects of hormone therapy in estrogen receptor-positive primary breast cancer. *Cancer Res.* **61**, 8452–8458 (2001).
25. Jiang, J. *et al.* Phosphorylation of estrogen receptor- α at Ser167 is indicative of longer disease-free and overall survival in breast cancer patients. *Clin. Cancer Res.* **13**, 5769–5776 (2007).
26. Fisher, B. *et al.* Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J. Natl Cancer Inst.* **97**, 1652–1662 (2005).
27. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).
28. Johnson, W. E. *et al.* Model-based analysis of tiling-arrays for ChIP-chip. *Proc. Natl Acad. Sci. USA* **103**, 12457–12462 (2006).
29. Carroll, J. S. *et al.* A pure estrogen antagonist inhibits cyclin E–Cdk2 activity in MCF-7 breast cancer cells and induces accumulation of p130–E2F4 complexes characteristic of quiescence. *J. Biol. Chem.* **275**, 38221–38229 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Carroll, S.-F. Chin, I. Mills, C. Massie and P. Edwards for discussions; J. Mitchell for histology work; M. Eldridge, S. Vowler and B. Adryan for bioinformatic help; S. Buttiglieri for the p-TARGET-PAX2 construct; J. Eeckhoutte for the pcDNA-AIB-1 construct; and M. Iddawela for the gift of Herceptin. This work was supported by grants from the National Institute of Diabetes and Digestive and Kidney Disease (R01DK074967 to M.B.) and the National Cancer Institute (P01CA8011105, and the DF/HCC Breast Cancer SPORE Grant to M.B.). We acknowledge support from the University of Cambridge, Cancer Research UK and Hutchison Whampoa Limited.

Author Contributions A.H. and J.S.C. conceived all experiments. A.H., K.A.H. and J.S.C. performed all experiments. T.R.G. and M.B. provided essential reagents and bioinformatics support for the ChIP-on-chip experiment. I.R.H. and R.I.N. provided essential cell line reagents. J.J. and S.A. collected primary breast cancer samples and performed immunohistochemistry, with help from W.J.H. A.H. and J.S.C. wrote the manuscript with advice from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.S.C. (jason.carroll@cancer.org.uk).

METHODS

Cell lines. The MCF-7, T47D, ZR75-1 and BT-474 human cell lines were grown as described previously²⁷. Tam-R cells were derived by long-term exposure to tamoxifen¹⁸ and grown under the same conditions as wild-type MCF-7 cells. Herceptin was added to the medium at a final concentration of 10 μ M.

ChIP-on-chip experiments. Genome-wide ER ChIP-on-chip experiments were performed in duplicate, as described previously², with the exception that the Affymetrix seven Genechip tiling array 2.0R set was used. Analysis of ER-binding sites were determined with MAT²⁸, with a false discovery rate of 5%. Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-TABM-563.

Motif enrichment. Analysis of motif enrichments was performed with the CEAS program (<http://ceas.cbi.pku.edu.cn/>). The PAX motif is represented with Weblogo (<http://weblogo.berkeley.edu/>).

Plasmids. PAX2 expression was from p-TARGET-PAX2 (a gift from S. Buttiglieri), AIB-1 expression was from a pcDNA-AIB-1 construct (a gift from J. Eeckhoutte) and SRC-1 expression was from pSG5-SRC-1.

RT-PCR. Cells were deprived of hormones as described previously²⁹. Total RNA was collected and RT-PCR was performed as described previously². Primer sequences are provided in Supplementary Data 15.

ChIP. ChIP experiments were performed as described previously³. Antibodies used were anti-ER α (HC-20), anti-AIB-1/RAC3 (M-397) and anti-HDAC-1 (sc-6298 and sc-6299) from Santa Cruz Biotechnologies, and anti-PolII (ab5408), anti-H3R17 dimethyl (ab8284), anti-PAX2 (ab23799), anti-SRC-1 (ab2859) and anti-N-CoR (ab24552) from Abcam. Primer sequences are provided in Supplementary Data 15.

siRNA. siRNA experiments were performed as described previously². The sequence of the siRNAs were as follows: PAX2 siRNA (sequence 1), 5'-GAA-GUCAAGUCGAGUCUAUUU-3' (sense) and 5'-AUAGACUCGACUUGACU-UUUU-3' (antisense); PAX2 siRNA (sequence 2), 5'-CAUCAGAGCACAU-CAAAUCUU-3' (sense) and 5'-GAUUUGAUGUGCUCUGAUGUU-3' (antisense) (Dharmacon, USA); siN-CoR Smartpool (Dharmacon), containing the sequences 5'-GAUCACAUCUGUCAAUUUU-3', 5'-GAACGUGGCUCUC-

AAAGUUU-3', 5'-GAAAGGAAUUCGACACUGAUU-3' and 5'-GCCCUGGG-AUUUAUGAUGAUU-3'.

Western blotting. Cells were deprived of hormones as described previously²⁹. Antibodies used were anti-ER α (Ab-10) from Neomarkers (Lab Vision, UK); anti-ERBB2 (ab16901), anti-PAX2 (ab38738), anti-SRC-1 (ab2859), anti-N-CoR (ab24552) and anti- β -actin (ab6276) from Abcam; and anti-AIB-1/RAC3 (M-397) from Santa Cruz Biotechnologies.

Cell counting. Cells were plated at equal confluence and grown in full DMEM medium. Cells were transfected as described previously² and cells were stimulated with 100 nM oestrogen or 1 μ M 4-hydroxytamoxifen for 24 h, or in time-course experiments for the periods given in the figure. Total cells were harvested for automated cell counting using the Z2 Coulter Particle Count Analyzer.

Chromosome conformation capture (CCC) assay. CCC was performed in accordance with published protocols³⁰. The chromatin was digested with *Pst*I, and the real-time primers used were 5'-GGAGCGGAAGTGATTCAGAG-3' (forward) and 5'-TTGCAGAGACCTCTGGGAGT-3' (reverse). The TaqMan probe was 6-carboxy-fluorescein-5'-AGAGCAGTTCTGCTCTTCGC-3'. A control reverse primer against another *Pst*I site was included, namely 5'-AGAGTCACCAAGCCTCTGCAT-3'.

Immunohistochemistry. One hundred and nine ER-positive breast cancer sections were collected and processed as described previously²⁵. Immunohistochemistry for PAX2 was performed on an automated BondMax Immunostainer (Leica) with anti-PAX2 antibody (ab38738; Abcam) at a dilution of 1:100. AIB-1 immunohistochemistry was performed with anti-AIB-1 antibody (611105; Transduction Laboratories) at 1:200. Immunohistochemistry for ERBB2 was described previously²⁵. Examples of PAX2-positive and PAX2-negative stained samples are provided in Supplementary Data 16.

Statistical analyses. Analysis was performed in SPSS V16.0 for Mac (SPSS Inc.). Kaplan-Meier plots were constructed to display the data, and analysis was conducted with Cox regression. Time to relapse was taken as the outcome; binary variables for AIB-1 and PAX2 and the interaction between AIB-1 and PAX2 were included as predictor variables.

30. Hagege, H. *et al.* Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature Protocols* **2**, 1722–1733 (2007).

CORRIGENDUM

[doi:10.1038/nature07835](https://doi.org/10.1038/nature07835)**Regulation of *ERBB2* by oestrogen receptor–PAX2 determines response to tamoxifen**

Antoni Hurtado, Kelly A. Holmes, Timothy R. Geistlinger,
Iain R. Hutcheson, Robert I. Nicholson, Myles Brown, Jie Jiang,
William J. Howat, Simak Ali & Jason S. Carroll

Nature 456, 663–666 (2008)

In the Methods Summary, online-only Methods and the legend to Supplementary Fig. 16, the antibody catalogue number for the PAX2 antibody was incorrectly listed as ab38738. The correct antibody used for PAX2 immunohistochemistry was ab23799 (Abcam).

Role for perinuclear chromosome tethering in maintenance of genome stability

Karim Mekhail^{1,2}, Jan Seebacher², Steven P. Gygi² & Danesh Moazed^{1,2}

Repetitive DNA sequences, which constitute half the genome in some organisms, often undergo homologous recombination. This can instigate genomic instability resulting from a gain or loss of DNA¹. Assembly of DNA into silent chromatin is generally thought to serve as a mechanism ensuring repeat stability by limiting access to the recombination machinery². Consistent with this notion is the observation, in the budding yeast *Saccharomyces cerevisiae*, that stability of the highly repetitive ribosomal DNA (rDNA) sequences requires a Sir2-containing chromatin silencing complex that also inhibits transcription from foreign promoters and transposons inserted within the repeats by a process called rDNA silencing^{2–5}. Here we describe a protein network that stabilizes rDNA repeats of budding yeast by means of interactions between rDNA-associated silencing proteins and two proteins of the inner nuclear membrane (INM). Deletion of either the INM or silencing proteins reduces perinuclear rDNA positioning, disrupts the nucleolus–nucleoplasm boundary, induces the formation of recombination foci, and destabilizes the repeats. In addition, artificial targeting of rDNA repeats to the INM suppresses the instability observed in cells lacking an rDNA-associated silencing protein that is typically required for peripheral tethering of the repeats. Moreover, in contrast to Sir2 and its associated nucleolar factors, the INM proteins are not required for rDNA silencing, indicating that Sir2-dependent silencing is not sufficient to inhibit recombination within the rDNA locus. These findings demonstrate a role for INM proteins in the perinuclear localization of chromosomes and show that tethering to the nuclear periphery is required for the stability of rDNA repeats. The INM proteins studied here are conserved and have been implicated in chromosome organization in metazoans^{6,7}. Our results therefore reveal an ancient mechanism in which interactions between INM proteins and chromosomal proteins ensure genome stability.

Eukaryotic rDNA is tandemly repeated from about 100 to more than 10,000 times⁸. rDNA repeats provide the foundation for at least one ribosome-manufacturing compartment, the nucleolus. The budding yeast *Saccharomyces cerevisiae* has 100–200 rDNA units tandemly arranged on chromosome XII (Chr. XII) and forming one nucleolus (Fig. 1a, b)⁸. In addition to harbouring rRNA-coding DNA sequences, each unit contains intergenic spacers (IGS1 and IGS2) that promote repeat integrity (Fig. 1a)^{9–11}. Recruitment of nucleolar protein complexes RENT (regulator of nucleolar silencing and telophase exit; composed of Cdc14, Net1/Cfi1 and Sir2) and Cohibin (mitotic monopolin proteins Lrs4 and Csm1) to IGS1 suppresses unequal recombination at the repeats^{3,12–16}. This suppression is seemingly linked to the ability of these complexes to induce rDNA silencing, which involves chromatin changes preventing RNA polymerase II (Pol II)-driven transcription within IGSs of rDNA^{4,5,16–19}.

Purification of Cohibin suggested an association with INM proteins of unknown function¹⁶. To gain insight into the possible role of

this association in nucleolar organization, we purified native Cohibin and INM proteins by tandem affinity purification (TAP). The TAP-tagged proteins are functional *in vivo* (ref. 16, and below). We detected purified complexes by silver staining and total protein mixtures were analysed by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS). Negative controls were untagged cells. Purification of Lrs4 and Csm1 yielded peptides of INM proteins Heh1 (helix extension helix 1, also called Src1) and Nur1 (nuclear rim 1, Ydl089w) (Fig. 1c, d, and Supplementary Table 1, part A)¹⁶. Heh1, the orthologue of human Man1, is a member of a family of INM proteins containing a highly conserved LAP–Emerin–Man1 domain (LEM, also called HEH; Supplementary Fig. 2)^{20–22}. LEM-domain proteins are linked to multiple clinical conditions through emerging roles in fundamental cellular processes, including gene expression and chromatin organization^{6,7,20,21,23,24}. Little is known about Heh1 and Nur1 (ref. 20), which we define here as chromosome linkage INM proteins (CLIP). Purification of either INM protein yielded peptides for both Heh1 and Nur1 (Fig. 1e, f, and Supplementary Information, section A). Purification of Heh2, an Heh1 homologue (Supplementary Fig. 2)²⁰, did not yield peptides for CLIP or Cohibin proteins (Fig. 1e, f, Supplementary Fig. 3c and Supplementary Table 1). Moreover, TAP-tagged Heh1, Lrs4 and Csm1 immunoprecipitated together with Myc13-tagged Lrs4, Heh1 and Nur1, respectively (Supplementary Fig. 3d). Migration of Heh1 to 115 kDa, instead of the predicted 95 kDa, led us to identify multiple post-translational modifications of the protein and fluctuation of Heh1 levels over the cell cycle with peaks at interphase and mitosis (Supplementary Figs 3a, e, f and 4). These findings physically link rDNA-associated complexes to INM proteins.

Peripheral association of genes is linked to silent chromatin assembly, which seemingly stabilizes repeats by limiting access to recombination proteins^{2,7}. Thus, CLIP may assemble at IGS1 to cooperate with RENT and Cohibin to silence transcription and inhibit unequal rDNA recombination. We therefore monitored unequal sister-chromatid exchange (USCE) by measuring the rate of loss of an *ADE2* marker gene from rDNA repeats. Deletion of Sir2, Lrs4 or Csm1 increased USCE, as expected (Fig. 2a, b, and Supplementary Table 4)^{16,25}. USCE also increased after deletion of Heh1 or Nur1, but not of Heh2 (Fig. 2a, b, and Supplementary Table 4). *heh1Δ nur1Δ* cells showed additive USCE defects compared to single mutants, suggesting that INM proteins have partly overlapping functions at rDNA. Moreover, deletion of Heh1, Lrs4 or Csm1 exacerbated the effect of losing Sir2 (Fig. 2a and Supplementary Table 4)¹⁶, suggesting that Sir2 stabilizes rDNA by means of processes that are both dependent on and independent of CLIP/Cohibin. Because increases in USCE affect rDNA copy number on Chr. XII, we analysed its size by using a contour-clamped homogeneous electric field (CHEF). Chr. XII measured about 2.83 megabase pairs in wild-type cells (about 190 rDNA units), and chromosome smearing in *sir2Δ* cells was indicative of

¹Howard Hughes Medical Institute and ²Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.

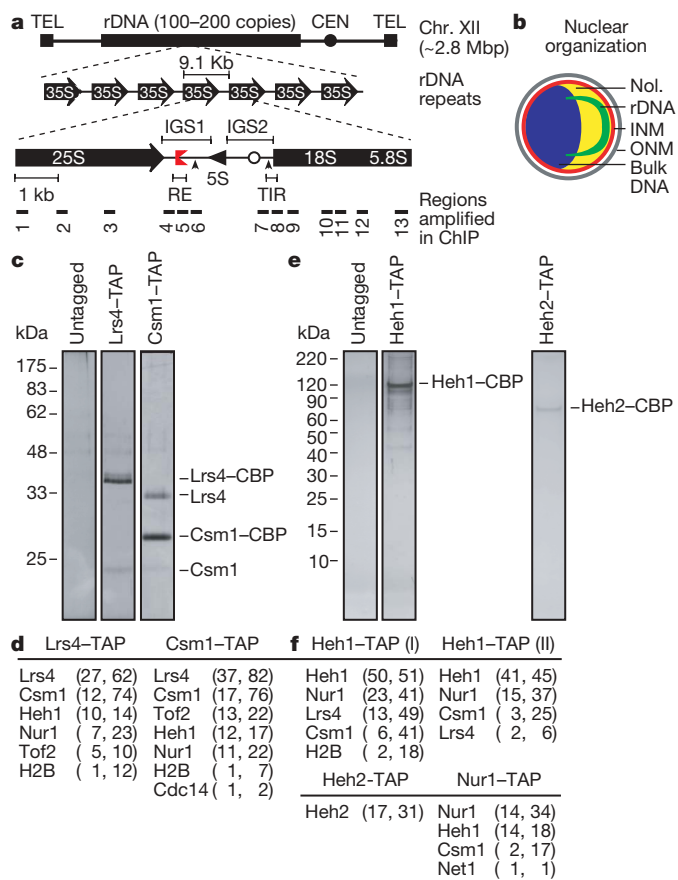


Figure 1 | Protein network extending from rDNA to the nuclear envelope. **a**, rDNA repeats on Chr. XII. Each unit yields a Pol I-transcribed 35S precursor rRNA (processed into 25S, 18S and 5.8S moieties) and a Pol III-transcribed 5S rRNA. CEN, centromere; TEL, telomere; IGS, intergenic spacer; RE, recombination enhancer; red symbol, replication fork block; TIR, Pol I transcription initiation region; open circle, DNA replication origin; Mbp, megabase pairs. Vertical arrowheads indicate insertion sites of *mURA3* reporter genes used in this study. **b**, Nuclear organization at the G_2/M cell-cycle stage. Nol., nucleolus; ONM, outer nuclear membrane. **c–f**, Purification of native complexes. **c**, **e**, Protein detection in silver-stained gels. TAP cleavage during protein purification leaves a calmodulin-binding protein (CBP) fragment. Results for purifications of Cohibin subunits (**c**) and INM proteins (**e**) are shown. **d**, **f**, LC-MS/MS analysis. Number of unique peptides followed by percentage coverage of the protein sequence is shown. Results for purifications of Cohibin subunits (**d**) and INM proteins (**f**) are shown. Full protein lists are given in Supplementary Table 1, part A, and spectral counts in Supplementary Tables 2 and 3.

severe changes in rDNA copy number (Fig. 2c and Supplementary Fig. 5a), as expected^{4,17}. Deletion of *Lrs4*, *Csm1*, *Heh1* or *Nur1* resulted in marked changes in rDNA copy-number averages and chromosome smearing patterns (Fig. 2c; described in Supplementary Information, section B). Together, these data suggest that the perinuclear protein network studied here is required for rDNA repeat stability.

We next studied the ability of cells to silence a Pol II-transcribed *mURA3* reporter gene positioned within IGS1 or IGS2 by assessing cellular growth on synthetic complete (SC) medium that either lacked uracil (–Ura; silencing disrupts growth) or was supplemented with 5-fluoro-orotic acid (+5FOA; silencing allows growth). Deletion of *Sir2*, *Lrs4* or *Csm1* disrupted IGS1 silencing (Fig. 2d), as expected¹⁶. However, in contrast to rDNA repeat stability, *Heh1* and *Nur1* were dispensable for silencing (Fig. 2d and Supplementary Fig. 5b), suggesting that silencing is insufficient for proper repeat-size regulation.

We next examined whether tethering rDNA repeats to INM proteins by means of Cohibin might limit recombination independently

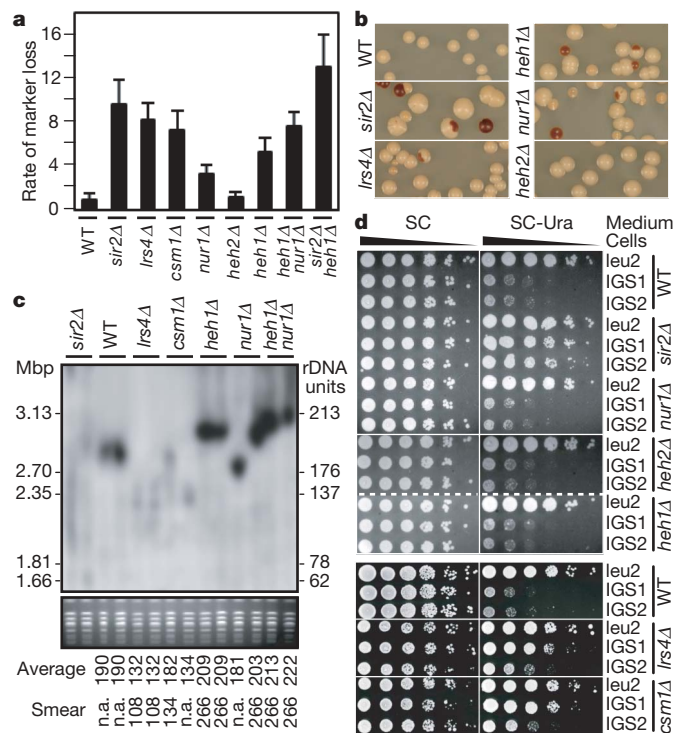


Figure 2 | Role of perinuclear protein network at rDNA repeats. **a**, **b**, Rates of *ADE2* marker loss (means \pm s.d.) relative to wild-type (WT) (**a**) and representative colonies (**b**) are shown. **c**, CHEF analysis of rDNA stability. Top: chromosomes resolved by CHEF were probed with IGS1 rDNA (Chr. XII). Bottom: ethidium bromide (EtBr) staining of Chr. IV and smaller chromosomes shows the quality of the preparation. Values corresponding to rDNA copy number averages and smear edges are indicated. Chr. XII size for *sir2Δ* was too heterogeneous for copy number to be estimated. **d**, Unlike *Sir2* and Cohibin, CLIP is dispensable for rDNA silencing. Tenfold serial dilutions of cells with the *mURA3* reporter gene inserted at IGS1/IGS2 (see Fig. 1a for locations) or outside rDNA at the *LEU2* locus are shown.

of silencing. Using immunofluorescence, we detected the functional green fluorescent protein (GFP)-tagged Net1 and Myc13-tagged *Heh1* (Supplementary Table 4)^{3,16}. Net1 associates with rDNA in the nucleolus throughout the cell cycle and recruits *Sir2* to IGS1 (ref. 3). However, enrichment of *Sir2* at rDNA in chromatin immunoprecipitation (ChIP) experiments is unaffected by deletion of Cohibin (Supplementary Figs 1 and 8c). To measure the limit of Net1–GFP internalization, the nucleus, delineated by peripheral *Heh1*–Myc13 signal, was divided into three concentric zones of equal area, zone I being the most peripheral²⁶. Cells were categorized according to whether the centre of the least peripheral Net1–GFP focus localized to zone I, II or III. Most wild-type cells showed peripheral Net1–GFP localization (zone I, 74%), and few contained central Net1–GFP staining (zone III, 2%) (Supplementary Fig. 6a). Deletion of *Lrs4* or *Csm1* drastically shifted Net1–GFP to zone II or III (Supplementary Fig. 6a) and expanded the volume occupied by Net1–GFP within nuclear space in three dimensions (Fig. 3a), suggesting that optimal perinuclear localization of the rDNA-associated Net1 requires Cohibin.

How rDNA is separated from the bulk of nuclear DNA is unknown (Fig. 1b). We tested whether the perinuclear network studied here affects this subnuclear separation. Nocodazole-arrested cells were analysed by fluorescence *in situ* hybridization (FISH) to reveal rDNA, and 4,6-diamidino-2-phenylindole dihydrochloride (DAPI) staining to reveal bulk nuclear DNA. Wild-type cells showed line-shaped rDNA spooling away from the DNA bulk towards the nuclear periphery (Fig. 3b; quantified in Supplementary Fig. 6b)²⁷. Deletion of *Lrs4*, *Csm1* or *Heh1*, but not *Heh2*, caused rDNA to adopt amorphous distributions often overlapping the DAPI signal, and a small percentage of cells showed two separable rDNA bodies (Fig. 3b

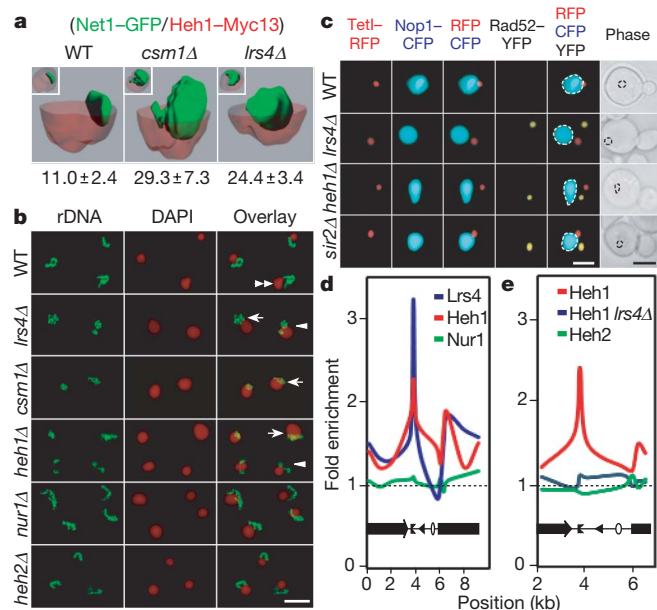


Figure 3 | Protein network tethers rDNA to the nuclear envelope.

a, Representative three-dimensional reconstructions of cells from double immunofluorescence analysis reveal the relative organization of Net1-GFP (green) and Heh1-Myc13 (red). Signal boundaries are shown with clipping planes for Heh1-Myc13. Percentages of green relative to full red volumes are shown (means \pm s.d.; $n = 5$). **b**, Nocodazole-arrested (G₂/M) cells were subjected to FISH to reveal rDNA and DAPI staining (pseudocoloured red) to detect bulk nuclear DNA. Cells with different rDNA morphologies are indicated by double arrowheads (spooled lines), arrows (amorphous) or triangles (two bodies) with quantifications in Supplementary Fig. 6b. Scale bar, 3 μ m. **c**, Live cells with Tet1-RFP-marked rDNA and expressing Rad52-YFP and nucleolar Nop1-CFP were imaged. Images depicting most observed phenotypes are shown. More images and quantifications are in Supplementary Fig. 6d–f. Scale bars, 1 μ m (white) and 3 μ m (black). **d**, **e**, Relative fold enrichment of indicated TAP-tagged proteins are shown. rDNA organization schematics are shown on graphs. kb, kilobases. Gels for **d** and **e** are shown in Supplementary Fig. 7a and Supplementary Fig. 7d, respectively. Detailed IGS1 ChIP is shown in Supplementary Fig. 7b, c.

and Supplementary Fig. 6b), which may have reflected a severe loss of interactions between rDNA repeats on Chr. XII sister chromatids. Deletion of Nur1 caused smaller changes in rDNA morphology, which appeared less condensed (Fig. 3b; $65 \pm 6\%$ of cells (mean \pm s.d.)). Disorganization of rDNA was also observed in asyn-chronous cells (Supplementary Fig. 6c).

We next studied the localization of a specific site within rDNA repeats in live cells harbouring a *tetO* array at rDNA repeats and expressing *tetO*-binding TetI-RFP (TetI fused to red fluorescent protein) and the nucleolar Nop1-CFP (Nop1 protein fused to cyan fluorescent protein)²⁸. TetI-RFP localized inside or at the periphery of the nucleolus in most wild-type cells (93%; Fig. 3c and Supplementary Fig. 6d, e). Deletion of *Lrs4* or *Heh1* shifted TetI-RFP outside the nucleolus or to its periphery (Fig. 3c and Supplementary Fig. 6d, e). *sir2Δ* cells showed less severe mislocalization of TetI-RFP (Fig. 3c and Supplementary Fig. 6d, e). Deletion of *Lrs4*, *Heh1* or *Sir2* also induced the formation of extranucleolar DNA repair centres, as marked by clustering of the yellow fluorescent protein (YFP)-tagged Rad52 recombination protein Rad52-YFP (Fig. 3c and Supplementary Fig. 6d, f). Fewer *sir2Δ* cells showed Rad52 foci than did *heh1Δ* or *lrs4Δ* cells (Supplementary Fig. 6f). This is in contrast to USCE in *sir2Δ* cells, which is higher than that in *heh1Δ* or *lrs4Δ* cells (Fig. 2a), suggesting that more recombinations in *sir2Δ* cells are unequal crossovers. Alternatively, a higher incidence of Rad52-YFP foci in cells lacking *Heh1* or *Lrs4* might suggest that these proteins stabilize several genetic loci. Most Rad52-YFP foci (61–65%) did not overlap TetI-RFP signal in *lrs4Δ*, *heh1Δ* or *sir2Δ* cells

(Fig. 3c and Supplementary Fig. 6d, f), probably reflecting the occurrence of one or few repair events per rDNA array and their distance from *tetO* sequences. Although we cannot exclude the possibility that the *tetO* array contributes to rDNA mislocalization in *lrs4Δ*, *heh1Δ* or *sir2Δ* cells, disruption of rDNA organization in *lrs4Δ* and *heh1Δ* cells lacking *tetO* sites, as revealed by FISH and immunofluorescence (Fig. 3a, b, and Supplementary Fig. 6a–c), argues against this possibility and suggests that the perinuclear complexes studied here help to stabilize wild-type rDNA repeats. Together, these results suggest that Heh1 and Lrs4, and to a smaller extent Sir2, are required for the sequestration of rDNA in the peripherally located nucleolus, and show that loss of sequestration is correlated with increased repeat instability (Fig. 2) and Rad52 recombination foci.

To further analyse CLIP–Cohibin links, we performed ChIP with a combination of dimethyl adipimidate and formaldehyde crosslinkers. We observed 2.85 ± 0.37 -fold and 2.35 ± 0.12 -fold enrichments for IGS1 sequences in Lrs4–TAP and Heh1–TAP immunoprecipitations, respectively (Fig. 3d, e, and Supplementary Fig. 7). We did not detect an enrichment by using Nur1–TAP, probably due to its low abundance or weaker association with Cohibin (Fig. 3d and Supplementary Figs 3a and 7a–c). More significantly, deletion of Lrs4 abolished the IGS1 enrichment of Heh1–TAP without affecting its levels (Fig. 3e and Supplementary Fig. 7d–f). In contrast, no enrichment was detected for Heh2–TAP (Fig. 3e and Supplementary Fig. 7d), an INM protein that neither interacts with Cohibin (Fig. 1) nor affects rDNA stability (Fig. 2), although it is expressed to similar levels as Heh1 (data not shown)²⁹. Together, these data indicate that CLIP–Cohibin-mediated tethering of rDNA repeats to the INM is required for repeat stability.

To determine whether perinuclear tethering suppresses recombination in the absence of Cohibin proteins, which are required for rDNA silencing and suppression of recombination, we created a strain in which rDNA was linked to Heh1 through Sir2. We fused *HEH1* and *SIR2* genes in *lrs4Δ* cells, creating a hybrid *HEH1-SIR2* gene (Fig. 4a). This yielded a fusion protein of the expected size (about 175 kDa) that was detectable by anti-Sir2 immunoblotting (Supplementary Fig. 8a). Fusion of Heh1 and Sir2 restored the separation of rDNA from bulk nuclear DNA in *lrs4Δ* cells (Fig. 4b), decreased unequal recombination (Fig. 4c and Supplementary Table 4) and increased homogeneity in the size of Chr. XII in cell populations (Fig. 4d and Supplementary Fig. 8b). Furthermore, ChIP revealed that Heh1-Sir2 became associated with rDNA at levels similar to those of Sir2 (Supplementary Fig. 8c). Moreover, Heh1-Sir2 did not rescue IGS1-specific increases in histone H3 acetylation, a marker for loss of silencing, caused by the deletion of *Lrs4* (Supplementary Fig. 8c). The inability of Heh1-Sir2 to restore rDNA stability fully may be due to other *Lrs4* functions, such as silencing or perhaps chromosome condensation, which suppress recombination at repeats. Attempts to fuse Heh1 with other perinuclear proteins, such as Tof2 or Ku70, did not yield viable cells (data not shown). Thus, tethering rDNA to the INM can promote repeat stability, at least partly independently of silencing.

Our results suggest that Sir2-dependent silencing alone cannot inhibit recombination within the repetitive rDNA locus and that INM-mediated perinuclear chromosome tethering ensures repeat stability (Fig. 4e and Supplementary Fig. 1). Extranucleolar Rad52 focus formation in *lrs4Δ*, *heh1Δ* or *sir2Δ* cells concurs with suggestions that although early rDNA recombination steps occur inside the nucleolus, Rad52 sumoylation and a high local concentration of the Smc5–Smc6 complex preclude Rad52 focus formation within nucleolar space²⁸. Thus, our findings suggest that rDNA repeats unleashed from the INM accumulate lesions that can better access the nucleoplasm, in which high concentrations of functional Rad52 promote DNA repair by homologous recombination. Therefore, perinuclear tethering probably sequesters repeats from recombination factors and may be required for Cohibin and RENT to align rDNA sister chromatids stably during replication to prevent unequal

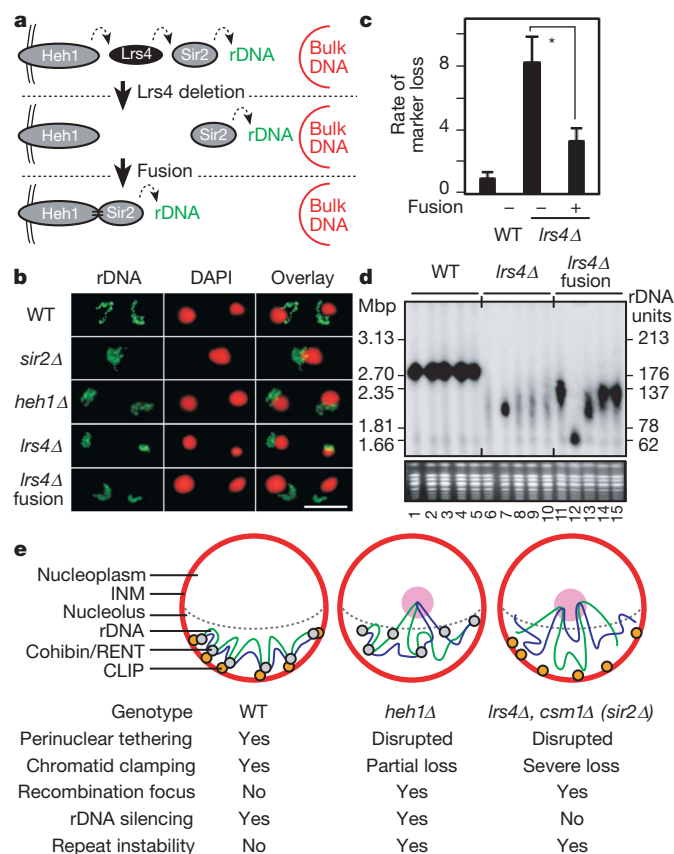


Figure 4 | Targeted perinuclear tethering promotes rDNA repeat stability. **a**, Fusion of Heh1 to Sir2 in *lrs4Δ* cells. Heh1 is shown embedded in the INM. **b**, FISH reveals that fusion of Heh1 and Sir2 restores the separation of rDNA signal from DAPI staining in nocodazole-arrested cells lacking Lrs4. Scale bar, 3 μm. **c**, Relative rates of *ADE2* marker loss (means \pm s.d.). Asterisk, $P < 0.0001$ for Student's *t*-test. **d**, CHEF analysis of rDNA stability. Top: chromosomes resolved by CHEF were probed with IGS1 rDNA (Chr. XII). Bottom: EtBr staining of Chr. IV and smaller chromosomes shows the quality of the preparation. **e**, Functional organization of the perinuclear molecular network tethering rDNA to the nuclear periphery. Repeat instability results from the loss of either the INM CLIP proteins or rDNA silencing complexes.

crossovers (Fig. 4e). Recombination between homologous repeats dispersed in the genome often instigates catastrophic chromosomal rearrangements. We expect that proteins studied here are members of perinuclear networks that control recombination at multiple loci to maintain genome stability.

METHODS SUMMARY

Standard co-immunoprecipitations^{3,14}, ChIP¹⁴, TAP purification¹⁶, immunofluorescence¹⁶, rDNA silencing¹⁶, USCE¹⁶, FISH²⁷ and live-cell²⁸ assays were performed as described previously.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 July; accepted 26 September 2008.

Published online 9 November 2008.

1. Szostak, J. W. & Wu, R. Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature* **284**, 426–430 (1980).
2. Moazed, D. Common themes in mechanisms of gene silencing. *Mol. Cell* **8**, 489–498 (2001).
3. Straight, A. F. *et al.* Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell* **97**, 245–256 (1999).
4. Bryk, M. *et al.* Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev.* **11**, 255–269 (1997).

5. Smith, J. S. & Boeke, J. D. An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev.* **11**, 241–254 (1997).
6. Capell, B. C. & Collins, F. S. Human laminopathies: nuclei gone genetically awry. *Nature Rev. Genet.* **7**, 940–952 (2006).
7. Reddy, K. L., Zullo, J. M., Bertolino, E. & Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**, 243–247 (2008).
8. Nomura, M. Ribosomal RNA genes, RNA polymerases, nucleolar structures, and synthesis of rRNA in the yeast *Saccharomyces cerevisiae*. *Cold Spring Harb. Symp. Quant. Biol.* **66**, 555–565 (2001).
9. Keil, R. L. & Roeder, G. S. Cis-acting, recombination-stimulating activity in a fragment of the ribosomal DNA of *S. cerevisiae*. *Cell* **39**, 377–386 (1984).
10. Brewer, B. J. & Fangman, W. L. A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* **55**, 637–643 (1988).
11. Kobayashi, T. & Ganley, A. R. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* **309**, 1581–1584 (2005).
12. Visintin, R., Hwang, E. S. & Amon, A. Ctf1 prevents premature exit from mitosis by anchoring Cdc14 phosphatase in the nucleolus. *Nature* **398**, 818–823 (1999).
13. Shou, W. *et al.* Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell* **97**, 233–244 (1999).
14. Huang, J. & Moazed, D. Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the replication fork block protein Fob1 in rDNA silencing. *Genes Dev.* **17**, 2162–2176 (2003).
15. Rabitsch, K. P. *et al.* Kinetochores recruitment of two nucleolar proteins is required for homolog segregation in meiosis I. *Dev. Cell* **4**, 535–548 (2003).
16. Huang, J. *et al.* Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *Genes Dev.* **20**, 2887–2901 (2006).
17. Gottlieb, S. & Esposito, R. E. A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. *Cell* **56**, 771–776 (1989).
18. Fritze, C. E., Verschueren, K., Strich, R. & Easton Esposito, R. Direct evidence for SIR2 modulation of chromatin structure in yeast rDNA. *EMBO J.* **16**, 6495–6509 (1997).
19. Smith, J. S., Caputo, E. & Boeke, J. D. A genetic screen for ribosomal DNA silencing defects identifies multiple DNA replication and chromatin-modulating factors. *Mol. Cell. Biol.* **19**, 3184–3197 (1999).
20. King, M. C., Lusk, C. P. & Blobel, G. Karyopherin-mediated import of integral inner nuclear membrane proteins. *Nature* **442**, 1003–1007 (2006).
21. Brachner, A., Reipert, S., Foisner, R. & Gotzmann, J. LEM2 is a novel MAN1-related inner nuclear membrane protein associated with A-type lamins. *J. Cell Sci.* **118**, 5797–5810 (2005).
22. Rodriguez-Navarro, S., Igual, J. C. & Perez-Ortin, J. E. SRC1: an intron-containing yeast gene involved in sister chromatid segregation. *Yeast* **19**, 43–54 (2002).
23. Hellemans, J. *et al.* Loss-of-function mutations in LEMD3 result in osteopetrosis, Buschke-Ollendorff syndrome and melorheostosis. *Nature Genet.* **36**, 1213–1218 (2004).
24. Bione, S. *et al.* Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. *Nature Genet.* **8**, 323–327 (1994).
25. Kaeberlein, M., McVey, M. & Guarente, L. The SIR2/3/4 complex and SIR2 alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes Dev.* **13**, 2570–2580 (1999).
26. Gartenberg, M. R., Neumann, F. R., Laroche, T., Blaszczyk, M. & Gasser, S. M. Sir-mediated repression can occur independently of chromosomal and subnuclear contexts. *Cell* **119**, 955–967 (2004).
27. Guacci, V., Hogan, E. & Koshland, D. Chromosome condensation and sister chromatid pairing in budding yeast. *J. Cell Biol.* **125**, 517–530 (1994).
28. Torres-Rosell, J. *et al.* The Smc5–Smc6 complex and SUMO modification of Rad52 regulates recombinational repair at the ribosomal gene locus. *Nature Cell Biol.* **9**, 923–931 (2007).
29. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Yip, T. Walz, J. Huang, M. Bühler, D. Koshland, A. Palazzo, D. E. Libuda, F. Winston, L. Vasiljeva, S. Buratowski, J. E. Warner, C. Anderson, G. A. Beltz, M. Lisby, T. Daniel, L. Ding and the Harvard NeuroDiscovery Optical Imaging Center for technical assistance or materials, and T. Rapoport, T. Iida, M. Motamedi, A. Johnson, M. Onishi, E. Gerace, S. Buker, M. Halic and members of the Moazed laboratory for helpful discussions and comments. This work was supported by the National Institutes of Health and the Howard Hughes Medical Institute (D.M.), and the Canadian Institutes of Health Research Institute of Aging (K.M.). D.M. is a scholar of the Leukemia and Lymphoma Society.

Author Contributions K.M. and D.M. designed experiments and wrote the paper. K.M. and J.S. performed LC-MS/MS analyses. K.M. performed the other experiments. S.P.G. provided mass spectrometry expertise and equipment.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.M. (danesh@hms.harvard.edu).

METHODS

Strains and materials. Endogenous genes were deleted or modified with carboxy-terminal epitope tags as described^{14,16}. Strains harbouring *mURA3* reporter genes were described¹⁴. For Heh1–Sir2 fusion, *HEH1* was amplified with its promoter from genomic DNA with primers KM11 (5′-GATAactagTTCTGCC-TGTAGAGAG-3′) and KM12 (5′-GATAgggcccCAAATATGGCAACTCG-GA-3′). *SpeI*–*ApaI*-digested products were ligated into pRS314, yielding a plasmid used as a template to amplify *HEH1* with the upstream *TRP1* gene with primers KM14 (5′-CATTCAAACCATTTTCCCTCATCGGCACATTAAAG-CTGGATGTCTGTATTATTTTAC-3′) and KM17 (5′-CGCTAGTCTTTG-ATACGGCGTATTTCATATGTGGGATGGTTATTTGTTTTCAGCGGAAT-3′) adding regions flanking endogenous *SIR2* start site. Cells lacking the endogenous *HEH1* open reading frame, transformed with PCR products and selected on –TRP medium were PCR/immunoblotting-screened. Antibodies: anti-Myc-9E10 (Covance), anti-actin (Millipore), HRP-conjugated anti-TAP or anti-Myc (Invitrogen), anti-digoxin (Jackson Laboratories), anti-Ack9/Ack14 H3 (Millipore), anti-CBP (Open Biosystems), anti-cyclin-B2 and anti-GFP (A. Rudner), Rhodamine-tagged goat anti-mouse (Jackson Laboratories), Alexa488-labelled goat anti-rabbit (Molecular Probes), FITC-conjugated goat anti-mouse (Jackson Laboratories), FITC-conjugated swine anti-goat (Invitrogen), anti-Sir2 (ref. 30).

Protein purifications. Standard assays were performed as described¹⁶. Purifications incorporating CHAPS were performed as described¹⁶ with modifications: (1) 1% CHAPS was added to lysis and TEV-cleavage buffers. (2) CHAPS (0.05%) was added to CAM binding and elution buffers. (3) Regarding purified mixtures, 10–50% was subjected to electrophoresis/silver staining and half was precipitated with trichloroacetic acid for spectrometric analysis.

Mass spectrometry. Trypsin-digested mixtures were subjected to LC–MS/MS³¹ and MS/MS spectral analysis³² as described (less than 1% false positive rate). Proteins in untaged controls were removed. Spectral counts semiquantitatively measuring the relative abundance of proteins are given in Supplementary Tables 2 and 3. Excised gel bands were minced, destained, dehydrated and trypsin-digested before extraction of digests. Modifications were identified with SEQUEST Sorcerer (Sage-N Research) allowing variable methionine oxidation, serine/threonine phosphorylation and lysine ubiquitylation. Only unambiguous phosphosites³³ are reported.

FISH. Experiments were conducted as described²⁷. rDNA probes were a gift from V. Guacci or were prepared from *Bgl*/III fragments from plasmids p362 and p363, which contain the 5′ and 3′ halves of an rDNA unit, respectively, using the BioNick (Invitrogen) and digoxigenin (Roche) labelling systems²⁷. Scoring was conducted at the microscope; representative images adjusted for contrast and colouring are shown.

Imaging. Images were collected with an Axiovert 200 microscope (Carl Zeiss) coupled to an EM-charge-coupled device digital camera (Hamamatsu Photonics) or an Eclipse 80i microscope (Nikon). The positions of GFP spots were determined as described²⁶. The outer circle was set to coordinates at which the red signal shows the largest decrease in intensity, moving centrally, as revealed by ImageJ (National Institutes of Health). Scoring and measurements were conducted with Metamorph (Molecular Devices) and representative images were adjusted for background using levels and contrast in Photoshop (Adobe). Other software programs handling data were Office (Microsoft) and FreeHand (Macromedia).

Three-dimensional reconstruction. An average of four images for each of about 15 Z sections were generated with a Zeiss LSM510 (Carl Zeiss) upright confocal microscope (Harvard NeuroDiscovery Optical Imaging Center). Respective settings for red and green signals were LP-560 and LP-505 emission, 543 and 488 nm excitation, 50 and 62 μm pinhole. Constant background corrections were performed in ImageJ. Reconstruction was performed with volume tools of Imapris (Bitplane).

ChIP. Standard ChIP (Supplementary Fig. 8) was conducted as described¹⁴. Modified ChIP (Fig. 3 and Supplementary Fig. 7) was conducted as described^{14,16} with protein–protein crosslinkers added³⁴. Modifications are as follows. Yeast cultures (50 ml) were grown to $D_{600} \approx 0.8$. Cells were centrifuged, washed with

ice-cold PBS, suspended in 10 ml of ice-cold fresh protein–protein crosslinking solution (10 mM dimethyl adipimidate, 0.25% dimethylsulphoxide in PBS) and nutated at room temperature (22–24 °C) for 45 min. PBS-washed cells were resuspended in 50 ml of 1% formaldehyde in PBS for 11 h, then glycine was added to 125 mM. PBS-washed cells were resuspended in 400 μl of lysis buffer and subjected to bead-beating; procedures were continued as described^{14,16} except that RNase was added before proteinase K. Dilutions for immunoprecipitation and input DNA were 1:2 and 1:20,000, respectively. [α -³²P]dCTP-labelled and EtBr-stained products were quantified with Molecular Imager/QuantityOne (Bio-Rad) and Image ReaderLAS-3000/ImageGauge (Fuji), respectively.

USCE. Assays were performed essentially as described^{16,25}. Cells were grown to $D_{600} = 0.4$ – 0.8 , sonicated briefly, and spread (about 400 cells per plate) on thick plates (5 mg l^{−1} adenine). Incubation was at 30 °C for 5 days, at 4 °C for 2 days, then at room temperature for 3 days. Rates were obtained by dividing the number of half-sectored colonies by the total number of colonies excluding completely red colonies.

CHEF and Southern blotting. Experiments were conducted as described^{11,35,36}, with modifications. Saturated overnight culture (1 ml) was washed and suspended in 300 μl of EDTA/Tris (50 mM EDTA, 10 mM Tris-HCl pH 7.5). Zymolyase (2 μl; 20 μg μl^{−1} in 10 mM Na₂HPO₄ pH 7.5) and 500 μl of low-melting-point CHEF-quality agarose (1% in 125 μM EDTA pH 8.0, 42 °C; Bio-Rad) were added and the mixture was solidified in plug moulds at 4 °C. Plugs were incubated overnight in 1 ml of 10 μM Tris-HCl pH 7.5, 500 μM EDTA at 37 °C, then overnight in 1 ml of 2 mg ml^{−1} proteinase K in 10 μM Tris-HCl pH 7.5, 500 μM EDTA, 10 mg ml^{−1} N-lauroylsarcosine at 50 °C. Plugs were washed three times with EDTA/Tris (4 °C, 1 h per wash) and stored in 2 ml of EDTA/Tris (4 °C). Plugs were prepared at 5 × 3 × 1.5 mm³ and run (68 h, 3.0 V cm^{−1}, 300–900 s, 10 °C) on a 0.8% CHEF agarose gel in 0.5 × TBE/CHEF-DR-II (Bio-Rad). CHEF size markers were *Hansenula wingei* chromosomes (Bio-Rad). EtBr-stained gels were imaged then subjected to standard Southern blotting. Blots were crosslinked by ultraviolet and probed (65 °C, 16 h) with [³²P]dCTP-labelled IGS1.

Whole cell protein preparation. Lysates were prepared by bead-beating¹⁶ except for those shown in Supplementary Fig. 8a; the latter were prepared by trichloroacetic acid-coupled lysis because Heh1–Sir2 was otherwise unstable. For this, 2 × 10⁷ cells grown to $D_{600} \approx 0.75$ were washed and suspended in 500 μl of ice-cold water. Sequential additions of 75 μl of alkali/2-mercaptoethanol (1.85 M NaOH, 1.065 M 2-mercaptoethanol) and 75 μl of 50% trichloroacetic acid solutions were each followed by a 10-min incubation on ice. After centrifugation (10,000 r.p.m. at 4 °C for 10 min), pellets were suspended in loading buffer (1 × standard loading buffer, 1.42 M 2-mercaptoethanol, 83.2 mM Tris-HCl pH 8.8) and boiled, and supernatants were saved.

α-Factor arrest. Cells grown to $D_{600} = 0.2$ were incubated for 3 h with α-factor (10 μg ml^{−1}). Cells were washed and resuspended in fresh medium; samples collected every 15 min were frozen in liquid nitrogen.

30. Moazed, D. & Johnson, D. A deubiquitinating enzyme interacts with SIR4 and regulates silencing in *S. cerevisiae*. *Cell* **86**, 667–677 (1996).
31. Haas, W. *et al.* Optimization and use of peptide mass measurement accuracy in shotgun proteomics. *Mol. Cell. Proteomics* **5**, 1326–1337 (2006).
32. Boker, S. M. *et al.* Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. *Nature Struct. Mol. Biol.* **14**, 200–207 (2007).
33. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature Biotechnol.* **24**, 1285–1292 (2006).
34. Kurdistan, S. K. & Grunstein, M. *In vivo* protein–protein and protein–DNA crosslinking for genomewide binding microarray. *Methods* **31**, 90–95 (2003).
35. Oakes, M., Siddiqi, I., Vu, L., Aris, J. & Nomura, M. Transcription factor UAF, expansion and contraction of ribosomal DNA (rDNA) repeats, and RNA polymerase switch in transcription of yeast rDNA. *Mol. Cell. Biol.* **19**, 8559–8569 (1999).
36. Libuda, D. E. & Winston, F. Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*. *Nature* **443**, 1003–1007 (2006).

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

Suppose the doors to the lab featured a disclaimer for new graduate students saying: "Warning: years of hard work inside, with no guarantee of a career." Some might think twice before entering. Although many graduate students recognize the pitfalls of striving for a position in academia, others still remain idealistic about their prospects.

At least one laboratory's website now offers a measured dose of reality (www.biology.duke.edu/johnsenlab/advice.html). "Before you apply to this lab or any other," writes Sönke Johnsen on his group's website, "there are a few things to keep in mind." Johnsen, an associate biology professor at Duke University in Durham, North Carolina, stresses that graduate school in biology is not a sure path to success nor is it a certainty that students will achieve a career similar to that of their adviser. He underscores this argument by doing the maths. On average, a professor at a research university looks after 3 students at a time for about 5 years each, which equates to 18 students over a 30-year career. Given that the total number of academic positions has stayed roughly constant in recent years, these 18 people are, in effect, competing for one job.

Should more labs post such disclaimers? Quite possibly, given the positive response Johnsen's note has received from contributors at www.scienceblogs.com. No wise adviser would allow for false promises or false hope — although lab heads should also be sure to emphasize that PhD scientists, especially those open to non-academic jobs, are generally quite employable.

Johnsen goes on to offer another reality check, albeit one related more to quality of life than job openings. Make sure to pursue your passions, he writes, before committing to an intense five or six years in the lab. And make sure that you don't commit to a miserable, yet high-profile, lab on the assumption that the pay-off, half a dozen years down the line, will make it all worthwhile.

Once you've read and understood a disclaimer of this type, you're probably ready to walk through those doors.

Gene Russo is editor of *Naturejobs*.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Assistant European Manager:
Nils Moeller (4953)

Natureevents:
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com

Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Kerstin Vincze (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock
Naturejobs online production: Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718

Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

MOVERS

Exequiel Ezcurra, director, University of California Institute for Mexico and the United States, Riverside, California



2005-08: Provost, San Diego Natural History Museum, and director of the Biodiversity Research Center of the Californias, San Diego, California

2001-05: President, National Institute of Ecology, Ministry of the Environment and Natural Resources, Mexico City, Mexico

In a strange way, Argentina's political upheaval granted Exequiel Ezcurra opportunities to advance his scientific career. His communication skills made him a mainstay at the interface of politics and science and his new position, as the director of the University of California Institute for Mexico and the United States (UC MEXUS) in Riverside, California, will put those diplomatic skills to use.

When a military junta took power in 1976 and dissidents started 'disappearing', many scientists left Argentina or stayed abroad (see *Nature* 451, 494-496; 2008). Ezcurra's publication of several papers after graduation eased his path to the United Kingdom: the University College of North Wales in Bangor, invited him to do a master's degree and he stayed on to do a doctorate in plant ecology.

Ezcurra garnered funding from the non-profit Ford Foundation to pursue his PhD on desert ecology at the USA-Mexico border. He designed a biodiversity and vegetation survey of the Sonoran desert to help the local government create a protected area. But to his frustration, the incoming governor shelved the project. Ezcurra pursued his research career, while writing science books for the general public. In 1992, Mexican cabinet minister Luis Donaldo Colosio read one of these books, contacted Ezcurra and revived the decade-old survey to establish the protected area. "I learned that when you do good science to build a good case, the opportunity to use it will eventually arise," Ezcurra says.

Later, as research director of San Diego Natural History Museum in California, Ezcurra helped to make the museum's first IMAX movie, *Ocean Oasis*, a financial and critical success. "That was a life-changing experience for me because millions of people saw it — a much bigger audience than will read a scientific journal paper," he says.

In 2001, he became president of Mexico's new National Institute of Ecology. In addition to setting up successful rehabilitation and restoration projects, Ezcurra used the Internet to distribute all government environmental publications, including real-time air-quality monitoring data. His efforts instilled credibility in the nascent institute.

As the new director of UC MEXUS, Ezcurra steps into the shoes of founding director Arturo Gómez-Pompa, a renowned ecologist. Stanford ecologist Rodolfo Dirzo can't think of anyone better suited. "Exequiel knows the Mexican and US environments — ecologically, socially and politically — in a most impressive way," he says, lauding Ezcurra's universal world-view and scientific acumen. ■

Virginia Gewin

CAREERS & RECRUITMENT

Just cure it

Cancer researcher Brian Druker is adapting Nike's 'Just do it' slogan to 'Just cure it'. Oregon Health & Sciences University's (OHSU) Cancer Institute in Portland has received a \$100-million personal donation from Phil Knight, co-founder of the Portland-based footwear giant. The renamed OHSU Knight Cancer Institute has an ambitious agenda: to assemble a world-class personalized cancer-care research enterprise that incorporates genomics, bioinformatics, diagnostics and translational medicine.

"I'm hoping to move away from the twentieth-century paradigm of attacking cancer blindly by building the next generation of twenty-first-century cancer investigators," says Druker, who conducted pivotal clinical trials of Gleevec (imatinib), a drug used to treat chronic myeloid leukaemia. He says the new effort will include identifying critical targets for therapy, developing tools for early detection and more effectively using the immune system to make vaccines for cancers with viral aetiologies.

Druker will recruit up to five lead scientists and their staff immediately to fill existing laboratory space. But his aim is to find scientists not afraid to apply new technologies or to take risks. "I don't want my biases to lead

us only in one direction," he says.

Although he expects to hire a few senior scientists as well as up-and-coming researchers working on cancer breakthroughs, Druker says he first needs a leadership team in place. "My most pressing needs include a director of the basic cancer-research effort as well as people to direct our clinical translational research effort, particularly in breast and gastrointestinal cancers," he says. He will match clinicians with basic scientists to build the bridges needed for this multidisciplinary approach to succeed.

More clinicians will be needed later to treat patients in a phase-I clinical-trial programme. By creating a joint clinic-research enterprise, Druker will be able to funnel clinical revenue into the research institute.

Tom Beer, director of the institute's prostate cancer programme, says that the OHSU offers many career opportunities working on breast, gastrointestinal, genitourinary and lung cancers. "It is incredibly exciting to be part of an organization on a steep path of growth, which should be compelling to mid-career faculty members ready to stretch their wings," he says. In that regard, Beer says their biggest asset is Druker, with his energy, vision and expertise. ■

Virginia Gewin

POSTDOC JOURNAL

All work and no play

It was an ordinary day at the office. My nose was a frozen spike of ice, rain pattered on my leather hat, and I'd momentarily lost my focal group. A gelada genteelly passed gas behind me and I became aware of the soft babbling of 250 monkeys on the way to their sleeping site. There was a big smile on my face. What an office!

I feel more inspired by this work than any other employment I've undertaken. Perhaps the isolation is having an effect. This field site stands in sharp contrast to the site for my PhD, where there was a lively social scene. I enjoyed that time immensely. Dare I admit that I spent most of my days in the field there pondering the intricacies of relationships, social blunders and old-fashioned gossip?

Now — shockingly — I find myself thinking about work. Yes, I have times in which I daydream about my loved ones and the next holiday, but those are brief moments. My mind is so uncluttered that I am constantly evaluating experiments, trying to look at different angles of research questions, wondering which PDFs I need to find when I'm in civilization again.

At least during daylight hours, my thoughts instinctively run that way. Is this when I discover that my career is taking over my life? Send me someone — anyone! — and we can discuss the matter ... ■

Aliza le Roux is a postdoctoral fellow in animal behaviour at the University of Michigan.

Kidroid

There's no time like the present.

Shane Clark

Jenny met her first Kidroid when she was seven years old.

"Can I touch it, Daddy?" she asked her father. He was an unassuming man with brown eyes and a bowl-cut that made him look like he was five.

"Sure, Jenny. See the sign? It says you can." Daddy's attempt to teach Jenny how to say the funny letters on the sign fell flat, as she'd already tuned him out. Silly Daddy, she thought. Always talking and talking. Jenny crawled on top of the display box. She rubbed the Kidroid's fuzzy brown hair. She stuck her fingers in the Kidroid's fuzzy brown ears. She tickled its nose. The Kidroid looked at her, blinked once, twice. Then it went back to staring at nothing.

Daddy was talking again, and he was using the louder voice. It wasn't yet the loudest voice, so Jenny knew she had more time. Not like with Mommy, whose no was no. Jenny smacked the Kidroid's side, and punched the top of its head.

OK, loudest voice was here now.

"... Of all the ... ! *Get down here this instant, young lady!*"

Jenny hopped off the pedestal and tugged at her Daddy's pantlegs. "I'm sorry, Daddy, something just came over me!" Those were her new favourite words. She didn't really understand it, but "something just came over me" always made Daddy calm down.

"Just ... just don't do it again, Jenny. Those things aren't toys." He hugged her close. Yup. Like a magic spell.

But their power faded with time, as all childhood magic does. Jenny was nine now, and learning things like classification of mammals and United Nations history in school; she didn't climb on top of display cases anymore. One thing she still did was bug her father.

"Daddy! But ..." her face wrinkled up in a pouting frown.

"No!" He was harder on her now, and Jenny didn't quite understand it. It was something to do with Mom's accident.

"Aw, but ..." But there would be no buts, not anymore. His decision was final.

So, that night, she took matters into her own hands. She'd never done anything like this before: never crept out of her bed after midnight, never taken her bicycle out from the darkened garage. Her back tingled and shook the half-mile down to the shopping mall. She'd tied a flashlight onto the front

of her bike with duct tape for a headlight. It was her only ally. She tried not to think of wraiths and ghosts and wolves, but they thought themselves into her head anyway.

At last, she reached the supermarket doors. She tested the electronic eye — not on. She ran around, tried a few side doors. The third one worked. And then she was inside and staring into the face of a Kidroid.

"Hello," she said.

No answer.

"I'm Jenny."

No answer.

"Mom used to say it would be a bad thing if I got you, because you couldn't keep up with me. But you'd be with me all the time! What do you think?"

No answer.

On her thirteenth birthday, Jennifer asked for one present.

"A Kidroid?" asked Dad. He'd thought she'd grown out of that silly obsession ages ago. "I mean, sure, Jennifer, but don't you think you're a touch old for it?"

"Dad! It's my birthday, right?" Her tone demanded agreement. "And I should get what I want, not what you think it's proper for me to want, right?"

"Well, yes, I assume so ..."

"Good!" She hugged him close, nuzzled his hair. She'd grown so tall these past few years. It was hard for Dad to understand how quickly that'd happened.

She had a party for her friends the day before her birthday, a party for her family the day of. Her grandmother and grandfather and aunts and uncles gave her dothing, money and advice. After they left, Dad went to his closet to retrieve a final present.

"It wouldn't do to have you opening this while everyone else was around ..." he said. Jennifer flung her arms around him, because she knew what it was.

She opened it in her room. It was just like she'd remembered from the supermarket. Just like Sarah's was. Jennifer thought back to the shy, rich girl, the one she'd invited to birthday parties and hopscotch-jumpings when they were eight.

The Kidroid stood a foot tall: fuzzy brown hair, fuzzy brown ears. Tickle-able nose. Robotic gears wound up when she switched it on. Electricity coursed through its semisentient brain.

"Hello, child! And what might your name be?" The voice was sugary-sweet.

"Jenny!"

"And how old are you?"

"Thirteen!" It was imprinting on her, she knew, just like Sarah's had. Soon the Kidroid would be hers. *Hers*.

"That's pretty old. Before I become your friend, I have to ask you a question."

Wait. This wasn't part of the script. "Oh, just a question, dear, just a single question."

"Oh. Okay ..."

"Have you ever lost someone close to you?" The Kidroid's eyes pierced hers.

She was taken aback, but Jenny knew she had to answer. "My mother."

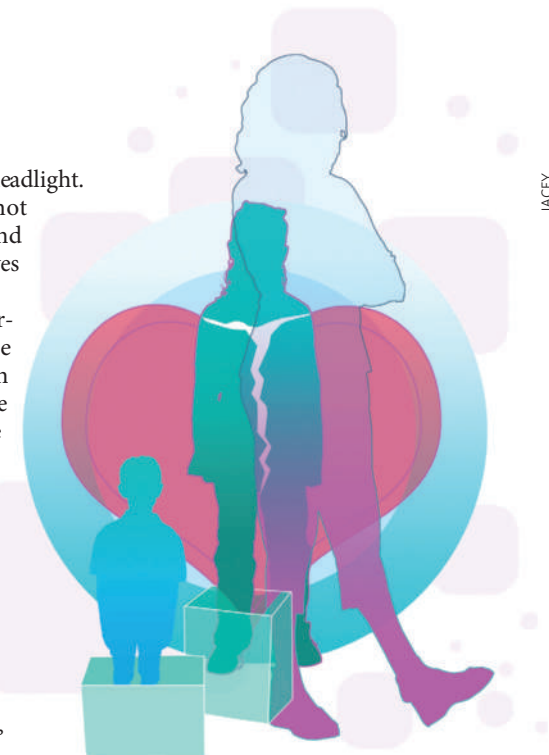
"And do you know what that means, to lose someone? To really *lose* someone?"

And then the tears came, came pouring out of her red eyes, because "my mother" was suddenly Mommy, was Rebecca Jane Porter, the lady with the flashing smile and the peanut-butter pancakes on Sunday mornings who she'd never, ever, ever see again.

The Kidroid blinked once, twice.

"Then it appears you're too old for me, after all." A tiny crackle. The smell of singed fur. The Kidroid's eyes closed for what Jennifer knew was the last time. She wordlessly scooped up the package and walked to the trash can. ■

Shane Clark is a biology student attending Georgetown University, working in an immunology lab and writing when he can. You can find him at <http://shaneclark.wordpress.com>.



JACEY